

Obtaining Nutritional Information for Dietary Assessment

Marc Zimmerman

A thesis
submitted in partial fulfillment of the
requirements for the degree of
Master of Science

University of Washington

2013

Committee:
Marian Neuhouser
Alanna Boynton

Program Authorized to Offer Degree:
Nutritional Sciences

©Copyright 2013

Marc Zimmerman

Table of Contents

1.	Introduction	1
1.1.	Specific Aims and Hypotheses.....	2
1.2.	Implications of Research	3
2.	Methods	3
2.1.	Recipe Definition	4
2.2.	Serving Sizes	5
2.3.	Nutrient Profile Definition	7
2.3.1.	Determining nutrient profiles for foods.....	7
2.3.1.1.	Aim 1: Investigation of Food Composition Data	8
2.3.2.	Generating a nutrient profile for a recipe.....	8
2.3.2.1.	Determining Consumption Information.....	9
2.3.2.2.	Assigning Weights Using Consumption Data	10
2.3.2.3.	Computing Nutrient Profiles	11
2.4.	Nutrient Database Analysis	12
2.4.1.	Aim 2: USDA Nutrient Database Sufficiency	12
2.4.2.	Aim 3: Evaluating Methods for Creating a Nutrient Database.....	13
3.	Results	14
3.1.	Recipe Definition	14
3.2.	Serving Sizes	16
3.3.	Determining Consumption Information.....	17
3.3.1.	Using Proxies	18
3.3.1.1.	Nutritional Variability within Recipe Components	18
3.3.1.2.	Amount of food harvested or produced for consumption.	19
3.3.1.3.	Market Share.....	20
3.3.1.4.	Consumer Dollars Spent.....	21
3.3.1.5.	Advertising	22
3.4.	Investigating USDA Database Sufficiency	22
3.5.	Investigating the Impact of Recipe Definition Strategy	24
4.	Discussion.....	26
4.1.	Aim 1 – Investigate Sources of Food Composition Data	26

4.2.	Aim 2 – Impact of relying solely on USDA SR-24.....	28
4.3.	Aim 3 – Impact of Recipe Definition Methodology.....	29
4.4.	Strengths and Limitations	30
5.	Summary	35

Tables

Table 1	Consumption of Fruit Juices by Children in the United States.....	10
Table 2	Weight computation for Foods in Line Item 1 Recipe.....	11
Table 3:	Market Shelf Survey – Sodas	17
Table 4:	Weight Assignments for Line Item 4, Flavored Waters.....	18
Table 5:	Amount Harvested of Vegetables Used to Define Line Item 18	19
Table 6:	Number of Servings of Vegetables Consumed in the United States	19
Table 7:	Weight Computations for Beverages Defining Line Item 5, Sodas.....	20
Table 8:	Final Weight Computations for Beverages Defining Line Item 5, Sodas.	21
Table 9:	Consumer Dollars Spent on Beverages Defining Line Item 7, Sports Drinks.....	21
Table 10:	Weight Computation for Beverages Defining Line Item 7, Sports Drinks.	22
Table 11:	Weight Computation for Beverages Defining Line Item 2, Fruit Drinks.	22
Table 12:	Relevant Foods Not Listed in USDA-SR24.....	23
Table 13:	Deviations in Nutrient Composition between USDA-sourced database and USDA-sourced database augmented with additional information.....	24
Table 14:	Deviations in Nutrient Composition Between Database Using Single-Food Recipes and Database using Multiple-Food Recipes.....	26

Figures

Figure 1:	Nutrition Labels for Common Sports Drinks.	15
-----------	---	----

Appendices

Appendix A:	Markets Surveyed for Market Shelf Analysis	38
Appendix B:	BSQ Line Items	38
Appendix C:	Computed serving sizes with sources	39
Appendix D:	Nutritional Databases.....	43

1. Introduction

Accurate dietary assessment is the foundation for determining nutritional status in individuals for clinical practice and in the population for public health.[1] Proper dietary assessment depends on the type of assessment tool used (e.g. food frequency questionnaire (FFQ), food record, etc.) as well as accurate food composition data.[2] The latter factor is less studied and often overlooked in dietary assessment studies, thus potentially leading to serious errors in dietary analysis.

The primary source for food composition data in the United States is the United States Department of Agriculture (USDA).[3, 4] The USDA developed and actively maintains the National Nutrient Database, which not only serves as the basis for most food and nutrition databases in the country, but is referenced in food policy, research and nutrition monitoring as well.[5] Specifically, this database currently provides information on 8,194 core food items and up to 146 food components. Each item in the database includes a description of the food, as well as its nutrient composition and common weights and serving sizes.

In order to provide this information, the USDA gathers data from both published and unpublished sources, including information compiled from scientific literature, the food industry, and government agencies. In addition, the USDA Agricultural Research Service (ARS) funds independent research efforts to assess the nutrient composition of several food products. These data are often based on laboratory analyses; however, nutrient information for some foods is computed using algorithms and/or recipes, rather than verified experimentally.[3]

Despite efforts to produce a comprehensive database, several data fields for existing food items are missing, and likely will not be populated for several years.[3] In addition, the USDA has minimal data on prepared and

packaged foods, including sweetened beverages, prepared snacks and chips, cookies, and other common beverages and snack foods. This information void is significant, because these foods constitute a substantial portion of the American diet. Estimates of market shelf space allocated to these foods – a proxy for consumption – is often more than 10 times greater than shelf space allotted to healthier foods such as fruits and vegetables.[6] Furthermore, these foods have been repeatedly linked to obesity, and so are a current focal point for nutritional research.[7] Unfortunately, this ongoing research could be hindered by the lack of valid assessment of the nutritional content of foods in the national nutrient database.

To address this problem, other sources of data must be used, such as those supplied by food manufacturers and market research groups. Efforts to obtain this information can be arduous due to its distributed and proprietary nature. However, it is nonetheless critical to improving the performance of dietary assessment tools and ultimately obtaining accurate dietary assessments.

1.1. Specific Aims and Hypotheses

The overarching goals of this study were to determine the adequacy of food composition data for sweetened beverages and snack foods, and to improve a previously validated beverage and snack food FFQ called the Beverage and Snack Food Questionnaire (BSQ),[8] by incorporating multiple sources of nutrient information into *recipe* definitions. The hypothesis was that a dietary assessment instrument with nutrient strings attached to each food item (or group of items) would provide a better nutritional status profile for the individuals or groups completing the instrument compared to the same instrument that assesses frequency of consumption only but is devoid of nutritional content data.

For the purposes of this study, a recipe is not a culinary recipe. Rather, it is a data set that is used to comprise a nutrient profile. A recipe is used to convert information from a FFQ into macronutrient and micronutrient values. This terminology is used by the Nutrition Data System for Research (NDSR), a dietary analysis tool at the Fred Hutchison Cancer Research Center (FHCRC) that was used in this study, and will be adhered to throughout this thesis.

The goal of this investigation was supported by the following aims:

1. To investigate the sources of food composition data for approximately 20 beverages and snack foods, including soft drinks, sports drinks, potato chips, cookies, and ice cream.
2. To populate a BSQ with nutrient values for several beverages and snack foods, comparing values from the USDA database alone to those from the USDA augmented with other market research.
3. To establish and evaluate the content validity of methods for creating and populating a nutrient database for the BSQ.

1.2. Implications of Research

If the source of food composition data is incomplete, then the accuracy of nutritional assessments based on this data could be compromised, potentially leading to underestimates of the effects of sweetened beverages and snack foods on nutritional status and energy balance. Understanding how food composition data and line item definitions impact the performance of FFQs may help researchers improve the accuracy of dietary assessment tools, and subsequently lead to a better understanding of the effects of foods on nutritional status.[9] This understanding will aid in the development of effective dietary recommendations for children and adults alike, and for the proper evaluation of dietary interventions.

2. Methods

This section describes the process of generating a food composition database, which was used to estimate the nutrient composition of BSQ line items. This process begins with defining recipes and serving sizes to represent the BSQ, and continues with computing nutrient profiles for these recipes.

After describing the generation of the food composition database, this section will then describe how this database was used to achieve the specific aims of this thesis, i.e. to assess the sufficiency of the USDA food database when analyzing the BSQ, and to evaluate the impact of using multiple foods to define a recipe, rather than defining recipes with a single food.

2.1. Recipe Definition

The first step of building a food composition database involved determining the specific foods that would be considered. To this end, recipes were developed to represent the line items of the BSQ. As described above, a recipe is a set of food composition data that is used to define the nutrient profile for a line item in a FFQ. For example, in order to assess the nutrient content of a serving of fruit, a researcher may subjectively use the nutrient profile of an apple. In this case, “apple” constitutes the recipe for a serving of fruit. In practice, recipes are often limited to a single food. However, associating a broad category of foods such as fruit with the nutrient profile of a single food can potentially compromise the precision of nutritional analyses. Therefore, this study attempted to improve the BSQ by incorporating multiple foods into recipe definitions.

Each recipe was restricted to a maximum of 4 foods. This restriction allowed each recipe to capture the nutritional variability encompassed by foods in each line item of the BSQ, while reasonably constraining the effort of obtaining food composition information. Furthermore, each recipe was constructed to encompass as diverse a set of foods as possible. For example, when determining the nutrient profile for *fruit*, 4 foods

were selected to comprise the recipe, incorporating a variety of tastes, nutrient profiles, etc. (e.g. apple, orange, banana, grape). Attempts were also made to include foods that are widely consumed, rather than uncommon or difficult to find foods that may be consumed by a small percentage of the population. For example, the recipe for *Ice creams and Milk Shakes*, line item 17 on the BSQ, includes *neapolitan* ice cream, rather than less common flavors such as *chocolate-cherry* or *cinnamon swirl*.

If foods and beverages referenced in a line item exhibited very similar nutritional profiles, the corresponding recipe was defined using fewer than 4 foods. In these situations, increasing the number of foods included in these recipes was not thought to impact nutrition analysis.

These recipes were discussed with and agreed to by a research team at the FHCRC.

2.2. Serving Sizes

Some FFQ's are semi-quantitative, and collect portion or serving size information from participants. However, the BSQ is non-quantitative. Therefore, in order to provide a nutrition analysis of foods in a recipe, serving sizes had to be determined. The USDA SR-24 [10] defines standard serving sizes for all food and beverage items in its database. When determining the serving size for a particular food in a recipe, this database was primarily considered. If the food existed in the database, the corresponding serving size in the database was used in the analysis.

The serving sizes for some products listed in the USDA SR-24 may not be consistent with serving sizes typically consumed. For example, the USDA database lists the standard serving size for fruit drinks (e.g. Snapple) as 8oz. However, fruit drink beverages are typically sold in larger volumes, and are likely consumed in a single

sitting. If the serving size listed in the USDA SR-24 was considered unrealistic by the research team, or if the product was not listed in the USDA SR-24, individual product labels were used to determine appropriate serving sizes.

Foods and beverages used in recipe definitions may be sold in a wide variety of portion sizes. For example, soft drinks are commonly available in 12-oz, 20-oz, and 1-liter containers. This variability complicates the process of selecting a single product label to use when determining an appropriate serving size. In these situations, a market shelf survey was conducted to determine the most popular serving sizes for the corresponding food or beverage.

A **market shelf survey** consisted of an estimation of supermarket shelf space devoted to a particular food or beverage, relative to similar foods or beverages. When conducting a market shelf survey, 3 types of food establishments were considered in order to develop a more comprehensive estimation of consumer dietary habits. Specifically, a convenience store, a natural foods retailer, and a major supermarket chain in the Seattle area were considered.

Each food retailer was surveyed in order to estimate the percentage of shelf space allotted to a particular type of food or beverage. These estimates were then averaged together to determine the mean shelf space allotted to a particular food or beverage. The most prevalent food and beverage products identified in the market shelf survey were used to determine the serving sizes for the foods or beverages in question. For example, if a market shelf survey found that, on average, 75% of shelf space for soda was occupied by 12-oz cans of soda, 12-oz was used as the serving size for soda. In addition to providing serving size information, market shelf surveys also served as a proxy for consumption data.[6] For example, if 25% of shelf space for

soda was occupied by diet soda, it was assumed that 25% of all soda consumed in the nation was diet soda. This consumption information was used in the development of nutrient profiles, and will be discussed later.

If a recipe contained a wide variety of foods and beverages available in multiple serving sizes, market shelf surveys were found to be impractical. For example, if a recipe contained chocolate chip cookies, apple pie, and vanilla ice cream, comparing shelf space allotment and serving sizes (e.g. 2 cookies, 1 piece of pie, ½ cup of ice cream) in an effort to determine a single serving size for the recipe is simply not feasible. In these cases, each food in the recipe was assigned a separate serving size, as defined by the Nutrition Data System for Research (NDSR). NDSR is a Windows-based dietary analysis program designed by the Nutrition Coordinating Center at the University of Minnesota. This tool incorporates a nutrient database derived from government-funded databases as well as other sources such as other food and nutrient databases, food manufacturers, and scientific journals.[11]

2.3. Nutrient Profile Definition

After defining recipes to represent the line items of the BSQ and determining appropriate serving sizes, each recipe's nutrient profile was determined. This process required the following 2 steps: 1) determining a nutrient profile for each food in a recipe, and 2) combining these profiles to create a single nutrient profile to represent the entire recipe.

2.3.1. Determining nutrient profiles for foods

In order to convert information from an FFQ into macronutrient and micronutrient values, it is essential to have food composition values.[8] For the purposes of this study, a nutrient profile consisted of the amount of

energy (kcal), added sugars, carbohydrate, protein, and fat contained in a single serving. These measures are the most commonly used in dietary assessment.

2.3.1.1. Aim 1: Investigation of Food Composition Data

To determine the composition of foods and beverages in each recipe, this study first considered government-funded databases (e.g., SR-24). However, if USDA databases were not able to provide this information, other sources of this information were surveyed. NDSR was the preferred alternative for this information. As described above, NDSR incorporates information from several sources to determine an appropriate serving size and nutrient profile for a given food or beverage. If nutrient composition values were neither available in SR-24 nor NDSR, other sources such as manufacturers' websites or product labels at local supermarkets were used to determine this information.

2.3.2. Generating a nutrient profile for a recipe

After determining a nutrient profile for each food in a recipe, the profiles were combined to yield a single nutrient profile to represent the recipe. In practice, this process often involves simply averaging the nutrient measures of a recipe's constituent foods. However, this strategy may yield an inaccurate assessment of nutrients in a recipe, particularly when some foods in a recipe are consumed significantly more widely than others. Therefore, a weighted average of the nutrient profiles in a recipe was computed. The weight assigned to a particular food was based on the consumption of that food relative to other foods in the recipe. For example, if a recipe for *fruits* was defined as "apples, oranges, bananas, grapes," the approximate frequency of consumption of each food in the population was determined (see 2.3.2.1, below). Each food was assigned a weight based on this consumption information. A weighted average of the nutrient profiles in a recipe was then computed. The use of food consumption information to assign weights to foods in a recipe is a novel

approach to nutritional analysis of FFQ's, and was intended to provide a more accurate nutritional assessment of the BSQ.

2.3.2.1. Determining Consumption Information

To determine consumption information, publications that utilize government-provided surveys such as the National Health and Nutrition Examination Survey (NHANES) were consulted. NHANES combines interviews and physical examinations to assess the nutritional status of both children and adults in the United States.[12]

If consumption of a particular food or beverage was not found in NHANES, an internet-based search for other sources of food consumption data was conducted. When possible, searches were limited to sources published within the past 6 years (i.e. 2007 – 2013). Both academic and private sources were surveyed to locate consumption information.

If no consumption data were available, then proxies for consumption were explored. Specifically, the following hierarchy was used to approximate consumption for each food:

- I. *Nutritional variability.* Variability between nutritional profiles for foods and beverages in a recipe was assessed. If these foods and beverages seemed similar for a nutritional standpoint, as determined by a research team at the FHCRC, then consumption of these items was ignored. Rather, each food or beverage in the recipe was weighted equally.
- II. *Amount of food harvested or produced for consumption.* It was assumed that 100% of food harvested or produced was consumed by the population. This approach to estimating consumption has been successfully used in previous work.[13, 14]
- III. *Market Share.* Industry reports often cite market shares of specific snack foods and beverages. It was assumed that market share was approximately equal to relative consumption of the particular food or beverage.

- IV. *Consumer dollars spent.* Like market shares, consumer dollars spent on foods and beverages are often cited by industry reports. It was assumed that consumer dollars spent on a food or beverage relative to similar foods and beverages approximated consumption of the food or beverage.

- V. *Dollars spent on advertising.* It was assumed that dollars spent on advertising for a food or beverage relative to the amount spent on similar foods and beverages approximated consumption of the particular food or beverage.

To approximate consumption for a food, the first proxy (i.e. nutritional variability) was considered. If this particular proxy could not be found, the next proxy was considered. This process continued until a proxy was found. This proxy was required to contain all foods in a recipe to facilitate more accurate comparisons.

2.3.2.2. Assigning Weights Using Consumption Data

After determining consumption information for each food and beverage in a recipe, weights were assigned accordingly using a relatively straightforward process. The amount of a food consumed relative to other foods in the recipe was computed, and this ratio was used as a weight. For example, line item 1, *Fruit Juices*, is defined as [*Apple Juice, Orange Juice, Grape Juice*]. To assign weights to these beverages, the consumption of each was identified, as shown in Table 1.

Fruit Juice	Percentage of Total Fruit Juice Consumed by Children[15, 16]
Apple Juice	35
Orange Juice	31
Grape Juice	25
Other	9

Table 1 Consumption of Fruit Juices by Children in the United States

Data sources typically include information for extraneous foods (e.g. *Other*, in the table above) that were not relevant to the recipe being considered. This information was excluded, and instead only pertinent foods

were considered. Specifically, the consumption of each food in the recipe relative to other foods in the recipe was computed as follows:

$$\text{Relative Consumption of Food } X = \frac{\text{Consumption of } X}{\text{Total Consumption of all Foods in Recipe}}$$

This process ensured that the total weight of each recipe was 1.00, which is a requirement of the nutrient analysis software used in this study. The weight computations for the recipe defining line item 1 are shown in Table 2.

Line Item	Recipe	Weight
1	Apple Juice	$\frac{35\%}{35\% + 31\% + 25\%} = 0.38$
	Orange Juice ¹	$\frac{31\%}{35\% + 31\% + 25\%} = 0.35$
	Grape Juice	$\frac{25\%}{35\% + 31\% + 25\%} = 0.27$

Table 2. Weight computation for Foods in Line Item 1 Recipe

2.3.2.3. Computing Nutrient Profiles

Nutrient profiles for each recipe were ultimately entered into the Nutrient Database Creator (NDC), which is a custom piece of software developed at the FHCRC. NDC allows the user to associate a recipe with each line item of a FFQ, and then assign a weight to each food in the recipe. These weights are applied to the nutrient profiles of their corresponding foods to determine the nutrient profile of the composite recipe, as described

¹ This particular recipe included two types of orange juice, Calcium-fortified and Vitamin-fortified. The weight was arbitrarily divided equally between each type (i.e. each of these beverages was assigned a weight of 0.175).

above. This nutrient composition database can then be used to automatically convert information in an FFQ into macronutrient and micronutrient information.

2.4. Nutrient Database Analysis

NDC was used to achieve Aims 2 and 3 of this thesis, specifically to investigate the sufficiency of the USDA nutrient database, and to evaluate methods for creating a nutrient database.

2.4.1. Aim 2: USDA Nutrient Database Sufficiency

To investigate the sufficiency of the USDA nutrient database, 2 databases were populated using NDC. Each database contained recipes and nutrient profiles to represent the line items of the BSQ. The first database included only foods with nutrient profiles available from the USDA food database. If a food was not listed in the USDA database, it was omitted from the recipe and not considered in the nutrient analysis of the line item. Furthermore, the weights assigned to the remaining foods in the recipe were recomputed to account for the omission. Specifically, the weight assigned to the omitted food was distributed equally among the remaining foods in the recipe such that the total weight for the recipe was 1.00. Relying solely on information provided by the USDA is the most common approach to nutritional analyses of FFQ's, and was considered the gold standard.

The second database included nutrient profiles for all foods in each recipe, independent of whether a food was listed in USDA SR-24. The profiles were available from NDSR. As described above, NDSR contains nutrient information extracted not only from the USDA food database, but from other information sources as well. This database was considered the experimental database.

The content validity of each database was then assessed in order to determine the extent to which recipes represented the nutritional domain of each line item. This assessment was completed by a team of experts at the FHCRC, including the creators of the BSQ.

After assuring the content validity of each database, NDC was used to provide a nutritional analysis of each line item. Specifically, the energy and macronutrient composition of each line item was assessed. Although NDC is able to provide a comprehensive analysis of recipes including micro- and macronutrients, consideration of all nutrients was beyond the scope of this study. Energy, protein, fat, and carbohydrate content are arguably the most common nutritional markers used in research and clinical practice. The databases were then compared on a line item by line item basis with respect to these metrics in order to illustrate the impact of relying solely on the USDA database to perform nutrient analyses of the BSQ.

Percent deviations were computed using the following formula:

$$\% \text{ deviation} = \frac{\text{Nutrient in recipe of experimental database (g)} - \text{Nutrient in recipe of standard database (g)}}{\text{Nutrient in recipe of standard database (g)}} \times 100$$

A positive deviation indicates that the recipe in the experimental database contains *more* of the nutrient marker (e.g. sugar) than the corresponding recipe in the standard database. Conversely, a negative deviation indicates that the recipe in the experimental database contains *less* of the nutrient marker.

2.4.2. Aim 3: Evaluating Methods for Creating a Nutrient Database

To establish the content validity of defining a recipe with multiple foods, 2 separate databases were again populated with nutrient profiles of line items in the BSQ. The first database employed the “standard” method of recipe definition (i.e. each recipe was defined using a single food).[17] In order to determine which food would be used to define a recipe, all foods used in the recipe were considered, and that with the highest

estimated consumption was selected. This food was assigned a weight of 1.00 both to comply with NDC, and because weights are proportions – as a result, they must sum to 1.00. The second database (i.e. the “experimental” database) used multiple foods to define each recipe. This information was provided to an expert committee to establish content validity of the experimental database.[18] Nutrient profiles for both the standard and experimental databases were provided by NDSR.

After establishing the validity of the experimental database, the impact of using multiple foods to define each recipe (i.e. using a more precise recipe definition) was assessed. For each database, NDC was used to perform a nutritional analysis of each line item. Results were compared with respect to energy, protein, fat, and carbohydrate content, and the percent deviation in each of these nutritional markers was computed for each line item as described above.

3. Results

3.1. Recipe Definition

The process of developing recipes was generally straightforward. However, minor difficulties were encountered. First, some line items encompassed foods with an extremely wide variety of nutritional profiles that was difficult to represent by a recipe of 4 items. For example, line item 13 assesses intake of candy, which can refer to several different foods including jelly beans, chocolates, peanut bars, etc. In these exceptional circumstances, four representative foods were selected based on discussions with a team of research experts at the FHCRC.

Conversely, two line items encompassed foods or beverages with very little nutritional differences among them. In these situations, fewer than 4 foods were selected to define the recipes. Increasing the number of

representative foods in these recipes did not appear to improve precision of nutritional analyses. For example, the most common sports drinks available for purchase contain nearly identical nutritional profiles. As shown in Figure 1, the nutritional profiles for 8-oz servings of Gatorade and Powerade are identical with respect to energy (50 kcal), sugar (14g), fat (0g), carbohydrate (14g), and protein (0g):

Nutrition Facts	
Serving Size 8 fl oz (240 mL)	
Servings Per Container 4	
Amount Per Serving	
Calories 50	
	% Daily Value*
Total Fat 0g	0%
Sodium 110mg	5%
Potassium 30mg	1%
Total Carbohydrate 14g	5%
Sugars 14g	
Protein 0g	
Not a significant source of Calories from Fat, Saturated Fat, Cholesterol, Dietary Fiber, Vitamin A, Vitamin C, Calcium, Iron.	
* Percent Daily Values are based on a 2,000 calorie diet.	

Gatorade Thirst Quencher

Nutrition Facts	
Serving Size 8 fl oz (240 mL)	
Servings Per Container 4	
Amount Per Serving	
Calories 50	
	% Daily Value*
Total Fat 0g	0%
Sodium 100mg	4%
Potassium 25mg	1%
Total Carbohydrate 14g	5%
Sugars 14g	
Protein 0g	
Vitamin B3 10%	Vitamin B6 10%
Vitamin B12 10%	Magnesium t
t Not a significant source of calories from fat, saturated fat, trans fat, cholesterol, dietary fiber, vitamin A, vitamin C, calcium and iron.	
*Percent Daily Values are based on a 2,000 calorie diet.	

PowerAde Fruit Punch Sports Drink

Figure 1: Nutrition Labels for Common Sports Drinks.

As a result of these similarities, two of the most common sports drinks on the market were selected to define the sports drink recipe (i.e. Gatorade and Powerade). A similar situation arose in the definition of the flavored water recipe (line item 4). Because of the similar nutritional profiles presented by various flavored waters, two of the most common (i.e. Propel and Glaceau) were selected to define the corresponding recipe.

A copy of recipe definitions for all nutritional databases constructed can be found in appendix C of this thesis.

3.2. Serving Sizes

The USDA SR-24 provided serving sizes for a vast majority of foods used in this analysis. However, other sources of information were needed to determine serving sizes for exceptional food items. For example, the USDA SR-24 does not include information for energy drinks. Therefore specific product labels were used to determine serving sizes for the Monster, RockStar, Full Throttle, and Red Bull energy drinks that define line item 3.

Generally, each food or beverage was assigned a single serving size. However, assigning a serving size to sodas and diet sodas in line items 5 and 6 was problematic. Although the USDA-SR24 defines a serving size of soda to be 12 oz, this definition did not seem consistent with quantities of sodas often consumed, given the wide variety of serving sizes for these beverages that are commonly available to the public. This variety affects nutritional analysis of the BSQ. For example, regular (non-diet) sodas are frequently sold in both 12 and 20-oz containers.² This variability can cause a nearly 67% deviation in the nutritional analysis of sodas used to define line item 6, sodas. Therefore, a market shelf survey was used to identify an appropriate serving size for these beverages.

As shown in Table 3, the market shelf survey confirmed a significant variability in serving sizes commonly available to the public:

² Although nutritional labels suggest (8 or 12-oz) servings, it was assumed that beverages sold in these volumes are typically consumed as a single serving.

Soft Drinks (volume)	Shelf Space Allotment ³		Average Shelf Space Allotment
	Supermarket	Convenient Store	
Regular Soda (12 oz)	77%	91%	84%
Regular Soda (20 oz)	24%	9%	16%
Regular Non-cola (12 oz)	77%	93%	85%
Regular Non-cola (20 oz)	24%	7%	15%
Diet Soda (12 oz)	77%	89%	83%
Diet Soda (20 oz)	24%	11%	17%
Diet Non-cola (12 oz)	77%	87%	82%
Diet Non-cola (20 oz)	24%	13%	18%

Table 3: Market Shelf Survey – Sodas

Therefore, rather than assign a single serving size for sodas, the two most common serving sizes were identified (i.e. 12- and 20-oz) and listed as separate beverages in the recipe definition for line item 6, sodas. The market shelf survey was also used to define serving sizes for beverages in the recipe for line item 5, diet sodas.

NDSR was used to determine serving sizes for foods such as *cake* in line item 15 that encompassed a wide variety of sizes and shapes, such that a market shelf survey was impractical.

Serving sizes for all foods used in recipe definitions, as well as the source used to determine each serving size, can be found in Appendix C of this thesis.

3.3. Determining Consumption Information

Initially, it was anticipated that government-provided surveys such as NHANES would readily describe how often the particular foods used in recipe definitions were consumed by a given demographic. However, for

³ The amount of soda and diet soda products available at the specialty market surveyed was negligible. Therefore, specialty markets were ignored in this survey, as they were not considered to contribute significantly to consumption of sodas and diet sodas.

several reasons including limited breadth of data provided and difficulty of data access and manipulation, these resources were impractical.

A single source could not be located that provided reliable information about the consumption of all foods listed in all recipes. Rather, several sources were investigated to gather this information. Although a few of these sources were affiliated with academic or non-profit institutions, the bulk of the consumption data available was provided by private companies. Access to most of these data required purchased subscriptions, and therefore was not pursued in this study.

Despite this limitation, explicit consumption data for foods comprising several recipes in this analysis was able to be located. Proxies for consumption were used to estimate consumption for remaining foods.

3.3.1. Using Proxies

3.3.1.1. Nutritional Variability within Recipe Components

The nutritional profiles of foods and beverages defining some recipes were roughly equivalent. In these cases, consumption data was not used to assign weights to recipe component. Rather, each component was simply weighted equally. This approach was used to assign weights to the recipe components of line item 4, Flavored Waters, as shown in Table 4.

Line Item	Recipe	Weight
4	Propel	0.50
	Glaceau Vitamin Water Essential	0.50

Table 4: Weight Assignments for Line Item 4, Flavored Waters

3.3.1.2. Amount of food harvested or produced for consumption.

If explicit consumption data could not be identified for a particular food, the amount of the food harvested or produced for consumption was investigated. This information was used to estimate the consumption of vegetables defining the recipe for line item 18.[13, 14]

Specifically, the vegetables used to define the recipe for line item 18, as well as the amount of each that was harvested for consumption, are shown in Table 5.

Recipe	Amount Harvested in 2011 (100,000 lbs)[13]
Tossed Salad with Dressing	49743
Frozen Peas	6042
Fresh Green Beans	16125
Canned Corn	65995

Table 5: Amount Harvested of Vegetables Used to Define Line Item 18

It was assumed that 100% of these harvested amounts were consumed. Based on this information, the number of servings of each vegetable consumed was computed. This computation was facilitated by NDSR, which provided estimates for the volume occupied by 1 lb of a particular vegetable.[19] Estimates for the number of consumed servings of each vegetable are shown in Table 6.

Vegetable	Volume in 1 lb	Serving Size	Number of Servings (1,000's)	Weight
Tossed Salad with Dressing	7.96 cups	1 cup	395.95	0.46
Frozen Peas	1.19 cups	½ cup	14.38	0.02
Fresh Green Beans	4 cups	½ cup	129.00	0.15
Canned Corn	2.45 cups	½ cup	323.38	0.37

Table 6: Number of Servings of Vegetables Consumed in the United States

Weights were then assigned by computing the number of servings of a particular vegetable consumed relative to that of other vegetables. For example, the weight assigned to Frozen Peas was computed as follows:

$$\text{Weight} = \frac{\text{Servings of Frozen Peas}}{\text{Total Servings of Vegetables}} = \frac{14.38}{395.95+14.38+129.00+323.38} = 0.02$$

3.3.1.3. Market Share

Another proxy for consumption utilized was market share. Various industry reports cited market shares of specific snack foods used in recipe definitions. For example, this information was used to assign weights to the beverages in line item 5, diet sodas. According to Beverage Digest, the top diet soda brands by carbonated beverage market share are Diet Coke (9.6%), Diet Pepsi (4.9%), Diet Mt. Dew (2.0%), and Diet Dr. Pepper (1.8%).^[20] Market shares of other diet soda brands were assumed to be negligible, and therefore not considered.

The recipe for line item 5 was defined by [Diet Cola, Diet Non-Cola]. Therefore, the top diet soda brands by market share were classified as either *Diet Colas* or *Diet Non-Colas*. The market share of each beverage type relative to the other was then computed, and weights were assigned accordingly. These weights are shown in Table 7.

Recipe	Related Soda Brands	Market Share	Weight
Diet Cola	Diet Coke	9.6%	$\frac{(9.6\% + 4.9\% + 1.8\%)}{(9.6\% + 4.9\% + 1.8\% + 2.0\%)} = 0.89$
	Diet Pepsi	4.9%	
	Diet Dr. Pepper	1.8%	
Diet Non-Cola	Diet Mt. Dew	2.0%	$\frac{2.0\%}{(9.6\% + 4.9\% + 1.8\% + 2.0\%)} = 0.11$

Table 7: Weight Computations for Beverages Defining Line Item 5, Sodas.

This particular line item presented a unique difficulty. As described above, multiple serving sizes were used in the definition of the Diet Sodas recipe. Therefore, the market share for each beverage type was distributed among these serving sizes based on results of the market shelf survey. The final weights are shown in Table 8.

Line Item	Recipe	Weight
-----------	--------	--------

5	Diet Cola – 12 oz	83% x 0.89 = 0.74
	Diet Cola – 20 oz	17% x 0.89 = 0.15
	Diet Non-Cola – 12 oz	82% x 0.11 = 0.09
	Diet Non-Cola – 20 oz	18% x 0.11 = 0.02

Table 8: Final Weight Computations for Beverages Defining Line Item 5, Sodas.

3.3.1.4. Consumer Dollars Spent

When considering some line items, such as line item 7, Sports Drinks, market share information was unavailable for one or more foods in a recipe.[21] However, market reports provided information related to consumer dollars spent on these foods and beverages. This information was then used to approximate relative consumption. For example, the beverages used to define the recipe for line item 7, as well as the consumer dollars spent on each beverage, are shown in Table 9.

Line item	Recipe	Consumer Dollars Spent in 2010 (\$1,000,000)[22]
7	Rockstar	400
	Red Bull	2,200
	Monster	1,500
	Full Throttle	110

Table 9: Consumer Dollars Spent on Beverages Defining Line Item 7, Sports Drinks.

It was assumed that consumer dollars spent on a particular sports drink relative to others roughly approximated the amount of the sports drink consumed relative to others. Therefore, the ratio of dollars spent on each sports drink in the recipe relative to others was computed, and this ratio was then used to assign a weight to each sports drink. The computed weights are shown in Table 10.

Recipe	Weight Computation
Rockstar	$\frac{400}{400 + 2,200 + 1,500 + 110} = 0.10$
Red Bull	$\frac{2,200}{400 + 2,200 + 1,500 + 110} = 0.52$
Monster	$\frac{1,500}{400 + 2,200 + 1,500 + 110} = 0.36$

Full Throttle	$\frac{110}{400 + 2,200 + 1,500 + 110} = 0.03$
---------------	--

Table 10: Weight Computation for Beverages Defining Line Item 7, Sports Drinks.

3.3.1.5. Advertising

When searching for consumption information for some items, explicit consumption data, market report data, or other information that could be used to estimate consumption information could not be obtained. In these cases, more obscure and potentially less accurate data sources were relied upon, including advertising dollars spent on products. Although advertising dollars spent on each recipe component may not correlate with actual consumption, the consumption estimates derived from advertising dollars spent seemed plausible, and so were used to assign weights to recipe components when necessary. This approach was used to determine weights for the recipe components of line item 2, Fruit Drinks. These weights are shown in Table 11.

Line Item	Recipe	Advertising Dollars Spent, 2011 (\$1,000)[23]	Weight Computation
2	Snapple	4,393	$\frac{4,393}{4,393 + 12,883 + 10,107} = 0.16$
	Capri Sun	12,883	$\frac{12,883}{4,393 + 12,883 + 10,107} = 0.47$
	Kool-Aid	10,107	$\frac{10,107}{4,393 + 12,883 + 10,107} = 0.37$

Table 11: Weight Computation for Beverages Defining Line Item 2, Fruit Drinks.

3.4. Investigating USDA Database Sufficiency

The USDA SR-24 proved to be an extensive source of nutrient information, and contained information for virtually every food contained in our recipes. In fact, the only foods not found in the USDA database are listed in Table 12.

Food	Line Item, Recipe
Snapple	2, Fruit Drinks
Kool Aid	2, Fruit Drinks
Glaceau Vitamin Water	4, Flavored waters
Goldfish crackers	12, Salty Snacks

Lifesavers	13, Candy
Jelly Beans	13, Candy
Large raised doughnut	14, Breakfast Pastries
Popsicle	16, Lowfat and nonfat frozen desserts

Table 12: Relevant Foods Not Listed in USDA-SR24.

As a result, the populated database containing information solely derived from the USDA was very similar to the experimental database containing information derived from the USDA and other sources. These databases are included in the appendix to this thesis.

Comparing BSQ nutritional analysis using only information from USDA-SR24 with analysis using this information augmented with other sources revealed minor differences (approximately 0-1%) for mean energy and macronutrients in 13 of 19 line items. However, as shown in Table 13, a deviation of more than 10% in at least 1 nutritional marker was observed in 4 line items. A deviation greater than 50% with respect to at least 1 nutritional marker was observed in 3 line items, specifically flavored waters, salty snacks, and frozen desserts. These deviations are solely attributable to the omission of specific foods from USDA SR-24.

Line Item	Line Item Name	Added Sugars (by Total Sugars) (g)	Energy (kcal)	Total Carbohydrate (g)	Total Fat (g)	Total Protein (g)
1	Orange juice, apple juice and other 100% juices	0	0	0	0	0
2	Fruit drinks (such as Snapple, flavored teas, Capri Sun and Kool Aid)	14.38%	12.05%	11.68%	5.24%	11.63%
3	Sport drinks (such as Gatorade or PowerAde); these drinks usually do not have caffeine	0	0	0	0	0
4	Flavored waters such as Propel or vitamin waters; these drinks usually do not have caffeine	100% ⁴	100% ⁴	100% ⁴	0	0
5	Diet soda or pop (include all kinds such as Diet Pepsi, Pepsi One, Diet Coke,	0	0	0	0	0

⁴ In the USDA-sourced database, all foods in the recipe for this line item contain 0g of this nutritional marker. However, in the experimental database, one or more foods in the recipe for this line item contain this nutritional marker.

	Diet 7-Up)					
6	Regular soda or pop (include all kinds such as Coke, Pepsi, 7-Up, Sprite, root beer)	0	0	0	0	0
7	Energy drinks (such as Rockstar, Red Bull, Monster and Full Throttle); these drinks usually have caffeine	0	0	0	0	0
8	1% or nonfat milk (sometimes called skim, fat-free, or low-fat milk; includes white and chocolate)	0	0	0	0	0
9	Regular or 2% milk (sometimes called whole, reduced fat, or 4% milk fat; includes white and chocolate)	0	0	0	0	0
10	Low-fat or non-fat potato chips, tortilla chips and corn chips (such as Baked Lays, Reduced-Fat Doritos, Fat-Free Pringles)	0	0	0	0	0
11	Regular potato chips, tortilla chips, corn chips and puffs (such as all flavors of Ruffles, Lays, Pringles, Doritos, Fritos, Cheetos)	0	0	0	0	0
12	Other salty snacks (like cheese nibs, Chex mix, gold fish crackers, Ritz Bitz)	65.29%	0.95%	2.96%	0.62%	0.40%
13	Candy, including chocolate, candy bars, jelly bellies, gummies and Lifesavers (do not include cookies)	0.51%	14.15%	1.64%	28.18%	28.09%
14	Doughnuts, pop tarts or other breakfast pastries	8.53%	1.33%	0.89%	5.03%	1.19%
15	Cookies, brownies, pies and cakes	0	0	0	0	0
16	Low or nonfat frozen desserts such as low fat ice cream, frozen yogurt, popsicles and sherbet	23.63%	32.05%	27.24%	40.67%	57.60%
17	Regular ice cream and milkshakes (include all flavors)	0	0	0	0	0
18	How often did you eat a serving of vegetables such as green salad, peas, green beans or corn? (do not count fried potatoes or French fries)	0	0	0	0	0
19	How often did you eat a serving of fruit such as a banana, apple or grapes? (do not count juices)	0	0	0	0	0

Table 13: Deviations in Nutrient Composition between USDA-sourced database and USDA-sourced database augmented with additional information.

3.5. Investigating the Impact of Recipe Definition Strategy

In order to assess the use of defining a recipe by using multiple foods, it was critical to first ensure that the recipes in the experimental database represented all facets of their corresponding line items at least to the extent of the recipes in the standard database, which were defined by a single food (i.e. the gold standard). To this end, content validity of the experimental database was established by subjective evaluation of an expert panel, as described above. According to the panel, including multiple foods in each recipe definition likely improves upon the extent to which each recipe represents the nutritional domain of a line item.

Nutritional analysis using single vs. multiple foods to define recipes revealed substantially more variability than observed in the previous analysis. As shown in Table 14, nutritional analysis of 18 of 19 line items revealed a deviation of more than 10% in at least 1 nutritional marker. Furthermore, a deviation greater than 50% in the assessment of at least 1 nutritional marker was observed in the nutritional analysis of 9 line items.

Line Item	Line Item Name	Added Sugars (by Total Sugars) (g)	Energy (kcal)	Total Carbohydrate (g)	Total Fat (g)	Total Protein (g)
1	Orange juice, apple juice and other 100% juices	0%	10.14%	8.98%	-1.91%	73.19%
2	Fruit drinks (such as Snapple, flavored teas, Capri Sun and Kool Aid)	14.38%	12.05%	11.68%	5.24%	-11.63%
3	Sport drinks (such as Gatorade or PowerAde); these drinks usually do not have caffeine	23.08%	23.08%	23.08%	0%	0%
4	Flavored waters such as Propel or vitamin waters; these drinks usually do not have caffeine	100.00% ⁵	100.00% ⁵	100.00% ⁵	0%	0%
5	Diet soda or pop (include all kinds such as Diet Pepsi, Pepsi One, Diet Coke, Diet 7-Up)	0%	-12.36%	-12.36%	-12.36%	-0.80%
6	Regular soda or pop (include all kinds such as Coke, Pepsi, 7-Up, Sprite, root beer)	0.44%	2.87%	2.26%	0.31%	-9.35%
7	Energy drinks (such as Rockstar, Red Bull, Monster and Full Throttle); these drinks usually	3.57%	-4.04%	-0.41%	-96.08%	-96.08%

⁵ In the baseline database (i.e. recipes defined by a single food), the food defining this line item contains 0g of this nutritional marker. However, in the experimental database (i.e. recipes defined by multiple foods), one or more foods defining this line item contain this nutritional marker.

	have caffeine					
8	1% or nonfat milk (sometimes called skim, fat-free, or low-fat milk; includes white and chocolate)	100.00% ⁵	11.36%	4.79%	83.10%	-0.27%
9	Regular or 2% milk (sometimes called whole, reduced fat, or 4% milk fat; includes white and chocolate)	100.00% ⁵	11.22%	5.30%	23.65%	-2.32%
10	Low-fat or non-fat potato chips, tortilla chips and corn chips (such as Baked Lays, Reduced-Fat Doritos, Fat-Free Pringles)	-25.00%	0.73%	-1.24%	8.05%	0.97%
11	Regular potato chips, tortilla chips, corn chips and puffs (such as all flavors of Ruffles, Lays, Pringles, Doritos, Fritos, Cheetos)	0%	-3.18%	5.46%	-13.21%	4.75%
12	Other salty snacks (like cheese nibs, Chex mix, gold fish crackers, Ritz Bitz)	100.00% ⁵	-6.65%	-2.27%	-12.23%	0.86%
13	Candy, including chocolate, candy bars, jelly bellies, gummies and Lifesavers (do not include cookies)	-84.24%	-27.28%	-23.90%	-27.81%	-32.85%
14	Doughnuts, pop tarts or other breakfast pastries	-27.26%	5.22%	-21.66%	41.39%	20.14%
15	Cookies, brownies, pies and cakes	-43.10%	-24.52%	-30.41%	-15.30%	-11.07%
16	Low or nonfat frozen desserts such as low fat ice cream, frozen yogurt, popsicles and sherbet	-13.65%	-38.64%	-25.36%	-75.13%	-65.12%
17	Regular ice cream and milkshakes (include all flavors)	50.84%	43.00%	47.53%	32.80%	57.32%
18	How often did you eat a serving of vegetables such as green salad, peas, green beans or corn? (do not count fried potatoes or French fries)	-122.22% ⁶	17.57%	50.05%	-70.01%	36.86%
19	How often did you eat a serving of fruit such as a banana, apple or grapes? (do not count juices)	0%	20.17%	20.04%	35.36%	11.22%

Table 14: Deviations in Nutrient Composition Between Database Using Single-Food Recipes and Database using Multiple-Food Recipes.⁷

4. Discussion

4.1. Aim 1 – Investigate Sources of Food Composition Data

The USDA SR-24 proved to be a surprisingly extensive source of nutrient information. With few exceptions, the database contained adequate food composition information to populate nutrient databases used to

⁶ Added sugar in the vegetable recipe is due to the inclusion of salad dressing. The recipe in the experimental database contains significantly less added sugar than the recipe in the baseline recipe on account of its inclusion of other vegetables in addition to salad with dressing.

⁷ All nutrient profiles for this analysis were provided by NDSR.

analyze the BSQ. The USDA SR-24 also provided thorough documentation with respect to the source of its nutrient analysis including upper and lower bounds on specific nutrient content, number of studies used to compute nutrient data, and the source of nutrient information (e.g. calculated, experimentally tested, etc.). Though not utilized in this study, this information may prove useful in future work addressing the derivation of nutrient information.

Nonetheless, the USDA SR-24 did not contain food composition information for all recipes used to define the BSQ. The USDA Food and Nutrient Database for Dietary Studies was also considered in order to determine food composition data for foods not listed in the USDA-SR24.[24] However, this database relies on information provided by the USDA SR-24, and so is subject to the same insufficiencies. As a result, obtaining this information required the investigation of alternate sources. This insufficiency highlights a shortcoming of the USDA SR-24, namely the absence of nutritional information for commonly consumed sweetened beverages and snack foods. These are precisely the foods that are implicated in the rise of the current obesity epidemic, and so availability of this information is critical to both clinicians and researchers in order to develop effective interventions.[25-33] There are certainly barriers to filling the current voids in the USDA database. Foods available to the public are constantly changing, and so it is not realistic to keep pace with these changes with complete accuracy. There is also a cost associated with augmenting the database. However, an argument can be made that those foods most commonly consumed should be prioritized.

Unfortunately, determining accurate consumption information for sweetened beverages and snack foods can be cumbersome, as experienced in this study. Although attempts were made to utilize results of the NHANES survey in order to assess consumption information, this resource was problematic. Understanding how to extract and analyze relevant datasets required completion of several tutorials which proved to be difficult to

use. Furthermore, the survey did not contain consumption information for several of the specific foods and beverages referenced in this study.

Several other sources were consulted to obtain consumption information. The most difficult part of this task was identifying a single source of information that described consumption (or a proxy for consumption) for every food in a recipe. Very often a source was identified that described consumption for some foods in a recipe, but not for others. For example, when investigating the consumption of foods referenced in the recipe for line item 11, *Regular potato chips, tortilla chips, corn chips and puffs*, a report was found that estimated consumption of potato chips and tortilla chips, but did not address consumption of corn chips and cheese puffs.[34] When considering foods in a recipe, obtaining consumption estimates from a single source was important, as these estimates were to be compared in order to determine relative weights. As a result, investigations continued until a single appropriate source was identified for each recipe. This experience was laborious, and underscored the fragmented state of available food consumption information.

4.2. Aim 2 – Impact of relying solely on USDA SR-24

When developing nutrient databases for this analysis, content validity was used to determine whether constructed databases adequately represented the nutritional domain of each line item in the BSQ. Content validity can be an important and appropriate tool to use given the qualitative nature of assessing whether a recipe adequately defines the nutrient domain of line item in a FFQ. Content validity is also easier to apply than more complicated validation strategies, is an accepted form of establishing validity, and can be more feasible to establish than other strategies that may require larger samples.[35]

Developing the nutritional database for the BSQ using only the USDA-SR24 was a straightforward process, as this resource provided composition information for most foods and beverages. However, composition information for some foods and beverages could not be found in the USDA-SR24. These foods were then omitted, resulting in an incomplete database.

Comparing this database with a complete database that utilized information provided by USDA-SR24 augmented with other sources presented several deviations. Although most deviations were considered minor (approximately 0-1%), deviations up to 100% were identified. Furthermore, in the USDA-sourced database, some recipes contained 0g of a particular nutritional marker, while the corresponding recipe in the experimental database contained a measurable amount of the same marker. For example, the recipe for line item 4, *Flavored Waters*, contains 0g of added sugar in the USDA-sourced database, but 15.69g in the experimental database. These types of deviations were reported as “100%” in the analysis.

These results are significant. In both clinical and research applications of the BSQ, over- or under-estimating nutrient composition to these extents can have adverse consequences including the misdiagnoses of nutrient excesses and deficiencies, and the development of inappropriate or ineffective nutritional interventions. Given these observations, the USDA database appears to contain informational gaps which may negatively impact nutritional analysis of the BSQ.

4.3. Aim 3 – Impact of Recipe Definition Methodology

When developing nutrient databases for FFQ's, nutrient profiles for line items are often defined by a single food. Incorporating several foods into these recipe definitions may increase the precision of nutrient databases by addressing the nutritional variability that exists between foods referenced in a particular line

item. In an effort to quantify the impact of using multiple foods to define a recipe versus the nominal procedure of using a single food to define a recipe, two databases were constructed to analyze the BSQ. The first database defined each line item using the nutrient profile of a single food -- specifically, the food that was estimated to be most widely consumed among those referenced by the line item. The second database defined each line item of the BSQ by computing a weighted average of nutrient profiles of up to 4 foods referenced by the line item.

Comparing these databases revealed discrepancies in a majority of line items. In fact, analyses of 9 out of 19 line items resulted in deviations of more than 50% in at least one nutritional marker. Based on these observations, using multiple foods to define a recipe appears to be a beneficial approach to nutritional analysis of the BSQ. This approach yields a theoretically more precise nutrition analysis that is significantly different from analysis using a nominal method of recipe definition based on a single food.

Determination of which recipe definition strategy was more precise was performed qualitatively in this study. Future work may investigate means of quantitatively establishing whether defining recipes with multiple foods yields a more precise nutrition assessment than definitions based on a single food. Quantitatively evaluating the precision of these approaches would likely require access to data from the target population (i.e. food records). Furthermore, recipe definitions in this study were restricted to a maximum of 4 foods. However, given these observations, it may be worth investigating the impact of using more than 4 foods to define recipes. This strategy can potentially improve the precision of BSQ analysis.

4.4. Strengths and Limitations

There are several strengths to this thesis that should be acknowledged. First, this study involved a thorough search for consumption information of specific beverages and snack foods by a particular demographic. Although several sources for consumption information are available, no single source was able to describe the consumption of all foods referenced in the analysis of the BSQ. Although sources of consumption information exist (e.g. NHANES), a PubMed search was unable to identify any other study that attempted to locate sources of consumption information, or to address the adequacy of consumption information available to academic and clinical researchers.

Another strength of this thesis is its use of multiple information sources to determine an appropriate serving size for each food and beverage considered. Accurate serving sizes are critical to dietary assessment, as they can significantly impact nutritional analyses.[36] While most developers of nutritional databases rely on a single source such as NHANES to determine serving size, this thesis used multiple sources as appropriate, including the USDA-SR24, market shelf surveys, and individual product labels.

This study was also unique in its attempt to assess the breadth of nutrient information provided by the USDA. Although dietary research often relies on the USDA to provide nutrient information, there is limited research that focuses on evaluating the completeness of this resource. In a recent study, Ahuja et al. described the evolution of the USDA databases to accommodate the changing American diet, but they do not offer any objective assessment of the database's ability to meet demands of the research community.[4] Pennington et al. also present a detailed discussion of the need for accurate food composition information.[1] While they reference deficiencies in the USDA nutrient database, they do not present any objective assessment of these deficiencies, nor do they examine the impact of these deficiencies on dietary assessment.

One study conducted by McCullough et al. assessed the consistency of information provided by several nutrient databases.[37] However, this study was primarily concerned with the accuracy of information in these databases, rather than the completeness of each. As well, this study also did not include an assessment of the USDA nutrient database. Furthermore, this study was conducted in the late 1990's. As a result, the authors' conclusions are somewhat inconsequential, given the perpetual modifications and improvements to which nutrition databases are subject.

While this thesis illustrated information gaps in the USDA nutrient database, Haytowitz et al.[38] addressed the issue of how to best fill these gaps. Given the cost associated with adding nutrient composition information to foods in the database, they recommend using consumption information to prioritize those foods that should be added to the USDA-SR. However, they do not address the availability of this consumption information, which has proven to be difficult to obtain for individual researchers. The federal government may be able to extract this information more readily (i.e. from NHANES).

This thesis used a weighted average of nutrient profiles, based on consumption of individual foods in a recipe, to improve the precision of the previously validated BSQ. Subar et al. considered several alternative methods for constructing nutrient databases and found that a data-driven method for constructing databases, such as that used here, yields a more precise nutritional assessment. Their study also found that algorithms based on means were more accurate and less error-prone than those based on medians.[39] These findings support the methodology used to construct the nutritional databases in this thesis.

Perhaps the biggest limitation of this study is the use of various types of information to estimate consumption information for foods used in recipe definitions. Ideally, consumption information would be used to determine appropriate weights for nutrition profiles of foods in a recipe. However, direct consumption

estimates for foods used to construct nutritional databases for the BSQ could not be located, requiring the investigation of various proxies for consumption. This investigation was arduous, which highlighted the general absence of reliable consumption information for sweetened beverages and snack foods such as those referenced in the BSQ.

Accuracy of the nutrition information provided by the USDA and other sources was assumed, and so no adjustments were made to nutrition information collected. As a result, the accuracy of nutritional analyses was limited by the accuracy of nutrition information provided by sources used. A brief survey of specific product labels indicated that accuracy of food composition information used in these analyses may be a legitimate concern. For example, as shown in Table 15, the fat content of a serving of soft-serve ice cream reported by the USDA was over 4 times the content reported by McDonald’s.

Food	Serving Size	Information Source	Energy (kcal)	Protein (g)	Fat (g)	Carbohydrate (g)
Soft-Serve Ice Cream – Vanilla	86g	USDA – SR24[10]	222	4.1	13	22.2
		McDonald’s[40]	139	4.1	3.7	22.1
Cake Doughnut – Plain	1 donut	USDA – SR24[10]	418	5.87	23.55	45.63
		Krispy-Kreme[41]	220	3	13	25

Table 15: Nutrient information provided by USDA and specific product labels.

This investigation also relied on USDA SR-24. However, USDA SR-25 was released to the public after the completion of this study, and so was not considered.[42] Information contained in this newer release may have filled information gaps identified in this investigation.

Determining foods that define a recipe is a subjective process. Therefore, a recipe may not fully represent the domain of the corresponding line item. This incomplete representation certainly affects the nutrient profile of a line item. However, efforts were made to minimize the impact of subjective recipe definitions by including only the most widely consumed foods, and by reviewing all recipe definitions with a team of experts at the FHCRC.

The content validity of nutrient databases constructed in this study is also limited by its subjectivity. Content validity is typically established by the subjective evaluation of an expert team. As a result, there is potential that the constructed databases do not fully represent the nutritional profile of foods referenced in the BSQ. However, the expert team used to establish content validity in this thesis included authors of the BSQ, which greatly reduces the potential for error in assuring the validity of these databases.

One source of information used to estimate consumption was the market shelf survey. This method of estimating consumption is arguably inaccurate, as these surveys were conducted only within Seattle-area grocery and convenience stores, and so results may not be generalizable to the rest of the nation. Furthermore, these surveys relied on subjective assessments of shelf space allotment, which is certainly prone to error. As a result, this information was used to estimate food or beverage consumption only in the absence of relevant published data. Another potential limitation of these surveys is that they discount consumption of foods and beverages from sources other than markets, including restaurants, cafeterias, vending machines, etc., which may impact serving size and consumption estimates for these foods and beverages.

Although use of multiple sources of information to estimate consumption may compromise the accuracy of nutrient analysis of the BSQ, it is indicative of a more significant problem facing nutritional research, namely the lack of accurate, consolidated information sources for food consumption in the United States.

5. Summary

In summary, this investigation revealed significant deviations in nutritional analyses the BSQ based on the source of food composition data. If the USDA-SR24 is considered the sole source of this data, the accuracy of nutritional assessments may be compromised. This imprecision may potentially lead to an over- or under-estimation of the impact of sweetened foods and beverages on nutritional status.

This investigation also identified significant variability resulting from the number of foods used to define each recipe. Incorporating multiple foods into recipe definitions theoretically yields a more precise nutrition analysis of the BSQ. Better understanding of how recipe definition impacts FFQ performance can potentially improve the accuracy of dietary assessment tools.

References

1. Pennington, J.A., et al., *Food composition data: the foundation of dietetic practice and research*, in *J Am Diet Assoc.* 2007: United States. p. 2105-13.
2. *Nutrition in the Prevention and Treatment of Disease*. 2nd ed, ed. A. Coulston and C. Boushey. 2008: Academic Press.
3. *USDA National Nutrient Database for Standard Reference, Release 25*. 2012.
4. Ahuja, J.K., et al., *USDA food and nutrient databases provide the infrastructure for food and nutrition research, policy, and practice*, in *J Nutr.* 2013: United States. p. 241S-9S.
5. USDA. *USDA Nutrient Data Laboratory*. [cited 2012 11/15/2012]; Available from: <http://fnic.nal.usda.gov/food-composition/usda-nutrient-data-laboratory>.
6. Farley, T.A., et al., *Measuring the food environment: shelf space of fruits, vegetables, and snack*. *J Urban Health*, 2009. **86**(5): p. 672-82.
7. Malik, V.S., et al., *Sugar-sweetened beverages and risk of metabolic syndrome and type 2 diabetes: a meta-analysis*, in *Diabetes Care*. 2010: United States. p. 2477-83.
8. Neuhauser, M.L., et al., *Development and validation of a beverage and snack questionnaire for use in evaluation of school nutrition policies*, in *J Am Diet Assoc.* 2009: United States. p. 1587-92.
9. Popkin, B.M., P.S. Haines, and A.M. Siega-riz, *Dietary patterns and trends in the United States: the UNC-CH approach*, in *Appetite*. 1999, 1999 Academic Press.: England. p. 8-14.
10. Service, A.R., *USDA National Nutrient Database for Standard Reference, Release 24*. 2011.
11. *NDSR User Manual - Appendix 22 - Sources of Nutrient Data*, U.o. Minnesota, Editor. 2012. p. A22.1 - 13.

12. Prevention, C.f.D.C.a. *About the National Health and Nutritional Examination Survey*. 2012 [cited 2013 2/1]; Available from: http://www.cdc.gov/nchs/nhanes/about_nhanes.htm.
13. Lucier, G. and L. Glaser, *Vegetables and Melons Outlook*, E.R. Service, Editor. 2011, USDA.
14. *Food Availability Documentaion*, E.R. Service, Editor. 2012, USDA.
15. Committee on, N., *Calcium Requirements of Infants, Children, and Adolescents*. Pediatrics, 1999. **104**(5): p. 1152-1157.
16. Dennison, B.A., *Fruit juice consumption by infants and children: a review*. J Am Coll Nutr, 1996. **15**(5 Suppl): p. 4S-11S.
17. Zimmerman, T. and S. Hull, *Development of food composition databases for food frequency questionnaires (FFQ)* Journal of Food Composition and Analysis, 2008. **21**: p. S20–S26.
18. Bannigan, K. and R. Watson, *Reliability and validity in a nutshell*. J Clin Nurs, 2009. **18**(23): p. 3237-43.
19. Minnesota, U.o. *NDSR*. [cited 2012 11/1/2012]; Description of NDSR tool developed by UMinn]. Available from: <http://www.ncc.umn.edu/products/ndsr.html>.
20. *Special Issue: U.S. Beverage Results for 2011*, in *Beverage Digest*. 2012. p. 1-2.
21. Heckman, M.A., K. Sherry, and E.G. De Mejia, *Energy Drinks: An Assessment of Their Market Size, Consumer Demographics, Ingredient Profile, Functionality, and Regulations in the United States*. BevNet. *The Top 15 Energy Drinks*. 2010 [cited 2013 March 28]; Available from: <http://www.energyfiend.com/the-15-top-energy-drink-brands>.
23. Harris, J., M. Schwarz, and K. Brownell, *Evaluating Sugary Drink Nutrition and Marketing to Youth 2011*, Yale Rudd Center for Food Policy and Obesity.
24. JKA, A., et al. *USDA Food and Nutrient Database for Dietary Studies, 5.0*. . 2012 [cited 2013 April 27].
25. Cavadini, C., A.M. Siega-Riz, and B.M. Popkin, *US adolescent food intake trends from 1965 to 1996*. West J Med, 2000. **173**(6): p. 378-83.
26. Nielsen, S.J., A.M. Siega-Riz, and B.M. Popkin, *Trends in energy intake in U.S. between 1977 and 1996: similar shifts seen across*. Obes Res, 2002. **10**(5): p. 370-8.
27. Mattes, R.D. and B.M. Popkin, *Nonnutritive sweetener consumption in humans: effects on appetite and food intake*. Am J Clin Nutr, 2009. **89**(1): p. 1-14.
28. Duffey, K.J. and B.M. Popkin, *Shifts in patterns and consumption of beverages between 1965 and 2002*. Obesity (Silver Spring), 2007. **15**(11): p. 2739-47.
29. Poti, J.M. and B.M. Popkin, *Trends in energy intake among US children by eating location and food source*. J Am Diet Assoc, 2011. **111**(8): p. 1156-64.
30. Lasater, G., C. Piernas, and B.M. Popkin, *Beverage patterns and trends among school-aged children in the US, 1989-2008*. Nutr J, 2011. **10**: p. 103.
31. Jahns, L., A.M. Siega-Riz, and B.M. Popkin, *The increasing prevalence of snacking among US children from 1977 to 1996*. J Pediatr, 2001. **138**(4): p. 493-8.
32. Bray, G.A., S.J. Nielsen, and B.M. Popkin, *Consumption of high-fructose corn syrup in beverages may play a role in the*. Am J Clin Nutr, 2004. **79**(4): p. 537-43.
33. Adair, L.S. and B.M. Popkin, *Are child eating patterns being transformed globally?* Obes Res, 2005. **13**(7): p. 1281-99.
34. *Nielsen's Shelf Stoppers/Spotlight: Snack Foods/Tortilla Chips*. 2012 [cited 2013 Feb 14]; Available from: <http://www.progressivegrocer.com/inprint/article/id2689/nielsen-s-shelf-stoppers-spotlight-snack-foods-tortilla-chips/>.
35. Sproule, C., *Rationale and Research Evidence Supporting the Use of Content Validation in Personnel Assessment* 2009.
36. McCabe-Sellers, B., *Advancing the art and science of dietary assessment through technology*, in *J Am Diet Assoc*. 2010: United States. p. 52-4.

37. McCullough, M.L., et al., *Comparison of 4 nutrient databases with chemical composition data from the Dietary Approaches to Stop Hypertension trial*. DASH Collaborative Research Group. J Am Diet Assoc, 1999. **99**(8 Suppl): p. S45-53.
38. Haytowitz, D., P. Pehrsson, and J. Holden, *The Identification of Key Foods for Food Composition Research*. JOURNAL OF FOOD COMPOSITION AND ANALYSIS, 2002. **15**: p. 183–194.
39. Subar, A.F., et al., *Evaluation of alternative approaches to assign nutrient values to food groups in food frequency questionnaires*. Am J Epidemiol, 2000. **152**(3): p. 279-86.
40. McDonald's. *Vanilla Ice Cream Cone - Nutrition*. [cited 2013 May 14]; Available from: http://www.mcdonalds.com/us/en/food/product_nutrition.dessertsshakes.178.vanilla-reduced-fat-ice-cream-cone.html.
41. *Doughnuts - Nutritional Information*. [cited 2013 May 14]; Available from: <http://www.krispykreme.com/nutri.pdf>.
42. USDA, *USDA National Nutrient Database for Standard Reference, Release 25*. 2012.
43. McCabe-Sellers, B. and ^{C. Chenard}, *Meeting the needs of US dietitians for food composition data*. Journal of Food Composition and Analysis, 2008. **21**(Supplement): p. S27–S34.
44. *Code of Federal Regulations, TITLE 21--FOOD AND DRUGS*. 2012.
45. Center, F.H.C.R. *General Nutrition Assessment (GNA) Documentation*. 2010 [cited 2013 April 7]; Available from: <http://sharedresources.fhcrc.org/services/food-frequency-questionnaires-ffg>.
46. Smiciklas-Wright, H., et al., *Foods commonly eaten in the United States, 1989-1991 and 1994-1996: are portion sizes changing?*, in J Am Diet Assoc. 2003: United States. p. 41-7.
47. Smiciklas-Wright, H., et al., *Foods Commonly Eaten in the United States: Quantities Consumed Per Eating Occasion and in a Day, 1994-1996*. . 2002, U.S. Department of Agriculture
48. *Popsicle - Did You Know?* 2012 [cited 2013 April 7, 2013]; Available from: <http://www.popsicle.com/article/detail/107644/popsicle-fun-facts>.
49. Dennison, B.A., et al., *Children's growth parameters vary by type of fruit juice consumed*. J Am Coll Nutr, 1999. **18**(4): p. 346-52.
50. Gould, B., *Understanding Dairy Markets*. 2006, Department of Agricultural and Applied Economics, University of Wisconsin, Madison.
51. *2009 Snack Sales*. 2009.
52. Group, T.G., *Potato Chips, Corn Chips, and Similar Snacks*. 2009.
53. *Nielsen's Shelf Stoppers/Spotlight: Snack Foods/Tortilla Chips*. 2012.
54. Seiz, K., *Branded pretzel category records outstanding growth*, in *Baking Management*. 2003. p. 16-18.
55. *RAB Instant Background Report for Snack Foods*, S.F. Association, Editor. 2002.
56. *America's 25 favorite candies*. 2009.
57. *Doughnut Statistics and Trends*. 2011.
58. Steinberg, S., *The Donut Book: The Whole Story in Words, Pictures & Outrageous Tales*. 2004: Storey Publishing, LLC.
59. *Top 10 Toaster Pastry Vendors*. 2005.
60. *Cake Statistics, 2010*. 2010, AIB International.
61. *Cake Statistics, 2011*. 2011, AIB International.
62. *Ice Cream and Frozen Desserts*. 2009.
63. *Ice Cream Production and Consumption Data*, C.D.I.C. Agriculture and Agri-Food Canada, Editor. 2006.
64. *Canned Peaches, Pears, and Fruit Mixtures: Conditions of Competition Between U.S. and Principal Foreign Supplier Industries*. 2007.
65. *Food Availability Documentation - Fruit (all uses)*. 2011.
66. *Fruit and Tree Nut Yearbook*. 2011.
67. *Jelly Beans*. 2006.

68. *Delivering Aspirational Growth*. 2007.

Appendix A: Markets Surveyed for Market Shelf Analysis

Supermarket	Address	Type
QFC	417 Broadway East, Seattle	General supermarket
Whole Foods Market	2210 Westlake Ave, Seattle	Specialty / Natural Foods Store
Eastlake Market	2244 Eastlake Ave E, Seattle	Convenience Store

Appendix B: BSQ Line Items

Order	Line Item Name
Q1	Orange juice, apple juice and other 100% juices
Q2	Fruit drinks (such as Snapple, flavored teas, Capri Sun and Kool Aid)
Q3	Sport drinks (such as Gatorade or PowerAde); these drinks usually do not have caffeine
Q4	Flavored waters such as Propel or vitamin waters; these drinks usually do not have caffeine
Q5	Diet soda or pop (include all kinds such as Diet Pepsi, Pepsi One, Diet Coke, Diet 7-Up)
Q6	Regular soda or pop (include all kinds such as Coke, Pepsi, 7-Up, Sprite, root beer)
Q7	Energy drinks (such as Rockstar, Red Bull, Monster and Full Throttle); these drinks usually have caffeine
Q8	1% or nonfat milk (sometimes called skim, fat-free, or low-fat milk; includes white and chocolate)
Q9	Regular or 2% milk (sometimes called whole, reduced fat, or 4% milk fat; includes white and chocolate)

Q10	Low-fat or non-fat potato chips, tortilla chips and corn chips (such as Baked Lays, Reduced-Fat Doritos, Fat-Free Pringles)
Q11	Regular potato chips, tortilla chips, corn chips and puffs (such as all flavors of Ruffles, Lay's, Pringles, Doritos, Fritos, Cheetos)
Q12	Other salty snacks (like cheese nibs, Chex mix, gold fish crackers, Ritz Bitz)
Q13	Candy, including chocolate, candy bars, jelly bellies, gummies and Lifesavers (do not include cookies)
Q14	Doughnuts, pop tarts or other breakfast pastries
Q15	Cookies, brownies, pies and cakes
Q16	Low or nonfat frozen desserts such as low fat ice cream, frozen yogurt, popsicles and sherbet
Q17	Regular ice cream and milkshakes (include all flavors)
Q18	How often did you eat a serving of vegetables such as green salad, peas, green beans or corn? (do not count fried potatoes or French fries)
Q19	How often did you eat a serving of fruit such as a banana, apple or grapes? (do not count juices)

Appendix C: Computed serving sizes with sources

Order	Recipe	Recipe Serving Size	Ref - Notes
Q1	Apple Juice	1 cup	FDA CFR-101.2 [44], GNA[45]
	Orange Juice - calcium fortified	1 cup	
	Orange Juice - vitamin fortified	1 cup	
	Grape Juice	1 cup	
Q2	Snapple	16 fl oz	<i>Snapple</i> (16 oz) product label
	Capri Sun	180 mL	<i>Capri Sun</i> pouch product

	Kool Aid	6.75 fl oz	label <i>Kool Aid (6.75 oz) bottle product label</i>
Q3	Gatorade	32 fl oz	<i>Gatorade (32 oz) product label</i>
	PowerAde	20 fl oz	<i>PowerAde (20 oz) product label</i>
Q4	Propel	24 fl oz	<i>Propel (24 oz) product label</i>
	Glaceau Vitamin Water Essential	20 fl oz	<i>Glaceau (20 oz) product label</i>
Q5	diet cola with caffeine - medium	12 fl oz	<i>Diet Coke (12 oz) product label</i>
	diet cola with caffeine - large	20 fl oz	<i>Diet Coke (20 oz) product label</i>
	diet non-cola - medium	12 fl oz	<i>Diet 7-up (12 oz) product label</i>
	diet non-cola - large	20 fl oz	<i>Diet 7-up (20 oz) product label</i>
Q6	cola with caffeine - medium	12 fl oz	<i>Coke (12 oz) product label</i>
	cola with caffeine - large non-cola - medium	20 fl oz 12 fl oz	<i>Coke (20 oz) product label 7-up (12 oz) product label</i>
	non-cola – large	20 fl oz	<i>7-up (20 oz) product label</i>
Q7	Rockstar	8 fl oz	<i>Rockstar (8 oz) product label</i>

	Red Bull	8 fl oz	<i>Red Bull</i> (8.3 oz) product label, GNA[45]
	Monster	8 fl oz	<i>Monster</i> (16 oz) product label
	Full Throttle	8 fl oz	<i>Full Throttle</i> (16 oz) product label
Q8	1% milk	8 fl oz	FDA CFR-101.2 [44], GNA[45]
	skim milk	8 fl oz	
	1% milk - chocolate	8 fl oz	
	skim milk - chocolate	8 fl oz	
Q9	2% milk	8 fl oz	FDA CFR-101.2 [44], GNA[45]
	whole milk	8 fl oz	
	2% milk - chocolate	8 fl oz	
	whole milk - chocolate	8 fl oz	
Q10	Lowfat Potato Chips - salted	28g	FDA CFR-101.2 [44], GNA[45]
	Lowfat-Baked Tortilla Chips – salted	28g	
Q11	Regular potato chips - salted	28g	FDA CFR-101.2 [44], GNA[45]
	Regular tortilla chips - salted	28g	
	Corn chips	28g	
	Cheese Puffs	28g	
Q12	Regular Cheese Crackers	30g	FDA CFR-101.2 [44]
	Chex Mix - traditional	29g	<i>Chex Mix</i> (8.75 oz) product label
	Gold fish crackers - original	30g	<i>Pepperidge Farm Goldfish Crackers</i> product label
	Regular Butter Crackers	30g	FDA CFR-101.2 [44]

Q13	Snickers candy bar	58g	GNA[45]
	Hershey's Milk Chocolate Bar with Almonds	41g	<i>Hershey</i> Milk Chocolate Bar label - 41g
	jelly beans - regular	40g	<i>Jelly Belly</i> Cocktail Shaker - Nutrition Label
	life savers	15g	<i>Life Savers</i> nutrition label serving size
Q14	regular doughnut - cake, plain	55g	FDA CFR-101.2 [44], GNA[45]
	regular doughnut - raised, glazed, plain	55g	
	regular doughnut - raised, plain	55g	
	regular pop tart	1 tart	
Q15	Cookies - chocolate chip, homemade from recipe, unknown fat	32g (2 medium cookies)	FDA CFR-101.2 [44], GNA[45], <i>Keebler</i> chocolate chip "Chocolate Lovers" cookie label (2 medium cookies / 32g), <i>Chips Ahoy</i> regular chocolate chip cookie label (33g)
	Regular chocolate brownies, plain, not frosted, unknown fat - commercially prepared, unknown egg content	40g	FDA CFR-101.2 [44], NDSR[19]
	Cake - Yellow, glazed, pudding type - added oil, unknown type of glaze, unknown egg, unknown fat in frosting (prepared at home)	80g (1 piece of cake)	FDA CFR-101.2 [44], GNA[45], NDSR[19]
	Apple Pie - double crust, commercial	1 piece of pie (162g)	GNA[45], NDSR[19]
Q16	Light ice cream - neapolitan, unknown fat %	1/2 cup (65g)	FDA CFR-101.2 [44], GNA[45], PSU report [46, 47], NDSR[19]
	Fat free sherbert	1/2 cup (85g)	FDA CFR-101.2 [44], GNA[45], PSU report [46, 47], NDSR[19]
	Regular popsicle	53g	Popsicle website[48]
	Low fat vanilla frozen yogurt	1/2 cup (100g)	FDA CFR-101.2 [44], GNA[45], PSU report [46, 47], NDSR[19]

Q17	Regular neapolitan ice cream	1/2 cup (65g)	FDA CFR-101.2[44], NDSR [19], PSU report[46, 47]
	Milkshake - hard vanilla ice cream	1 cup (228g)	FDA CFR-101.2[44], NDSR [19]
	Milkshake - soft vanilla ice cream	1 cup (186g)	FDA CFR-101.2[44], NDSR [19]
Q18	Tossed Salad with dressing	1 cup	GNA[45]
	Frozen green peas	1/2 cup	GNA[45]
	fresh green beans	1/2 cup	GNA[45]
	Yellow and white canned corn	1/2 cup	GNA[45]
Q19	Banana	1 medium piece	GNA[45]
	Apple	1 medium piece	GNA[45]
	Grapes	1/2 cup	GNA[45]
	canned fruit cocktail - light syrup	1/2 cup	GNA[45]

Appendix D: Nutritional Databases

Notes

- *Line Item Names* are taken directly from the BSQ.
- *Information Source* describes the source of nutrition information for a particular food, and was used to evaluate the sufficiency of the USDA database.
- *Serving size* describes the standard serving size for a particular food.
- *Frequency of Consumption* describes the frequency of consumption for a food, relative to the other foods in the recipe. If this information was unavailable, a proxy for consumption was used.
- *Weight* describes the associated weight that was assigned to a food's nutrient profile. It was based on the frequency of consumption of the food relative to the other foods in the line item definition.
- *Sugars, Energy, Protein, Fat, and CHO* describe the nutrient content of the line item. It was defined as a weighted average of the nutrients of foods in the corresponding recipe.

BSQ Nutrition Database Analysis – Information provided solely by USDA

Line Item	Line Item Name	Recipe	Info. Source (USDA or other)	Serving Size	Consumption Information or Proxy		Weight	Added Sugars (g)	Energy (kcal)	CHO (g)	Fat (g)	Protein (g)
Q1	100% Fruit Juices	Apple Juice	USDA	1 cup	% of total fruit juice consumption[49]	35%	0.38	0.000	126.960	30.788	0.316	0.925
		Orange Juice - calcium fortified	USDA	1 cup		31%	0.17					
		Orange Juice - vitamin fortified	USDA	1 cup			0.17					
		Grape Juice	USDA	1 cup		25%	0.27					
Q2	Fruit Drinks (e.g. Snapple)	Capri Sun	USDA	180 mL	\$1,000's spent on advertising, 2011[23]	12883	1.00	18.685	84.742	22.108	0.034	0.065
Q3	Sport drinks	Gatorade	USDA	32 fl oz	N/A	--	0.50	21.824	96.348	25.032	0.036	0.058
		PowerAde	USDA	20 fl oz		--	0.50					
Q4	Flavored waters (e.g. Propel)	Propel	USDA	24 fl oz	N/A	--	1.00	0.000	0.000	0.000	0.000	0.000

Q5	Diet soda or pop	diet cola with caffeine – medium	USDA	12 fl oz	2011 market share of carbonated beverage brands, 85% soft drink consumption is 12 fl oz, 15% 20 fl oz, based on market shelf survey[20]	16.3	0.76	0	6.938	1.006	0.105	0.426
		diet cola with caffeine – large	USDA	20 fl oz		16.3	0.13					
		diet non-cola – medium	USDA	12 fl oz		2.0	0.09					
		diet non-cola – large	USDA	20 fl oz		2.0	0.02					
Q6	Regular soda or pop	cola with caffeine – medium	USDA	12 fl oz	2011 market share of carbonated beverage brands, 85% soft drink consumption is 12 fl oz, 15% 20 fl oz, based on market shelf survey[20]	32.6	0.59	36.399	153.895	39.515	0.081	0.259
		cola with caffeine – large	USDA	20 fl oz		32.6	0.10					
		non-cola – medium	USDA	12 fl oz		14.3	0.26					
		non-cola – large	USDA	20 fl oz		14.3	0.05					
Q7	Energy	Rockstar	USDA	8 fl oz	2010 energy drink	400	0.10	26.626	110.391	27.807	0.104	0.325

	drinks (e.g. Red Bull)	Red Bull	USDA	8 fl oz	sales, \$1 millions[21, 22]	2,200	0.52					
		Monster	USDA	8 fl oz		1,500	0.36					
		Full Throttle	USDA	8 fl oz		110	0.03					
Q8	1% or nonfat milk	1% milk	USDA	8 fl oz	2006 milk sales, millions of lbs[50]	6,615	0.42	0.564	93.979	12.763	1.160	8.235
		skim milk	USDA	8 fl oz		8,320	0.53					
		1% milk - chocolate	USDA	8 fl oz		290	0.02					
		skim milk - chocolate	USDA	8 fl oz		365	0.02					
Q9	Regular or 2% milk	2% milk	USDA	8 fl oz	2006 milk sales, millions of lbs[50]	17,799	0.50	0.485	137.416	12.367	6.328	7.870
		whole milk	USDA	8 fl oz		16,410	0.46					
		2% milk - chocolate	USDA	8 fl oz		780	0.02					
		whole milk - chocolate	USDA	8 fl oz		719	0.02					
Q10	Low-fat or non-fat potato chips, tortilla chips and corn chips	Lowfat Potato Chips – salted	USDA	28g	2009 market share of potato chips (and related snacks) vs. tortilla chips, assume similar market shares among low-fat products[51-53]	64.1	0.80	1.498	116.250	22.718	2.173	1.868
		Lowfat-Baked Tortilla Chips –	USDA	28g		15.9	0.20					

		salted											
Q11	Regular potato chips, tortilla chips, corn chips and puffs	Regular potato chips – salted	USDA	28g	2009 sales, in \$1 millions[51-53]	7176.4	0.48	0	149.938	16.349	8.826	1.560	
		Regular tortilla chips - salted	USDA	28g		5579.8	0.38						
		Corn chips	USDA	28g		903.1	0.06						
		Cheese Puffs	USDA	28g		1217.4	0.08						
Q12	Other salty snacks (e.g. Chex Mix, Ritz Bitz)	Regular Cheese Crackers	USDA	30g	2002 cracker sales, split evenly between cheese crackers, gold fish, butter crackers, \$1 millions[54]	1,118	0.40	0.350	141.861	17.296	6.835	2.700	
		Chex Mix – traditional	USDA	29g		2002 sales, "savory snacks - variety packs", \$1 millions[55]	346						0.20
		Regular Butter Crackers	USDA	30g		2002 cracker sales, split evenly between cheese crackers, gold fish, butter crackers, \$1 millions[54]	1,118						0.40
Q13	Candy, including chocolate, candy bars,	Snickers candy bar	USDA	58g	2009 sales, \$1 millions[56]	1720.2	0.46	15.420	249.323	29.468	14.057	3.858	
		Hershey's	USDA	41g	2009 sales, \$1	1393.2	0.54						

	jelly bellies, gummies and Lifesavers	Milk Chocolate Bar with Almonds			millions[56]							
Q14	Doughnuts, pop tarts or other breakfast pastries	regular doughnut - cake, plain	USDA	55g	2005 doughnut sales, \$1 millions[57, 58]	242.67	0.2700	9.020	219.809	31.000	9.240	3.195
		regular doughnut - raised, glazed, plain	Other	55g	2005 doughnut sales, \$1 millions[57, 58]	242.67	0.2600					
		regular pop tart	USDA	1 tart	2005 toaster pastry sales[57, 59]	480.00	0.4700					
Q15	Cookies, brownies, pies and cakes	Cookies - chocolate chip, homemade from recipe, unknown fat	USDA	32g (2 medium cookies)	2010 sales of cookies / brownies / cakes / pies per store, \$1 millions [60, 61]	930	0.21	45.348	420.885	63.057	17.908	3.184
		Regular chocolate brownies, plain, not frosted, unknown fat - commercially prepared,	USDA	40g		138	0.03					

		unknown egg content										
		Cake - Yellow, glazed, pudding type - added oil, unknown type of glaze, unknown egg, unknown fat in frosting (prepared at home)	USDA	80g (1 piece of cake)		2714	0.61					
		Apple Pie - double crust, commercial	USDA	1 piece of pie (162g)		644	0.15					
Q16	Low or nonfat frozen desserts	Light ice cream - neapolitan, unknown fat %	USDA	1/2 cup (65g)	Amount of light ice cream / frozen yogurt / sherbert produced in 2009, 65g servings[62]	2,661	0.6200	12.892	130.299	22.725	2.949	3.468
		Fat free sherbert	USDA	1/2 cup (85g)	Amount of light ice cream / frozen yogurt / sherbert	282	0.1900					

					produced in 2009, 85g servings[62]							
		Low fat vanilla frozen yogurt	USDA	1/2 cup (100g)	Amount of light ice cream / frozen yogurt / sherbert produced in 2009, millions of 100g servings[62]	336	0.1900					
Q17	Regular ice cream and milkshakes	regular neapolitan ice cream	USDA	1/2 cup (65g)	Assume even distribution, unable to find milk shake consumption information[63]	0.34	0.34	18.002	243.173	31.610	10.804	5.568
		milkshake - hard vanilla ice cream	USDA	1 cup (228g)		0.33	0.33					
		milkshake - soft vanilla ice cream	USDA	1 cup (186g)		0.33	0.33					
Q18	Vegetables	Tossed Salad with dressing	USDA	1 cup	Number of 1 cup servings harvested for consumption[13]	395954.28	0.46	0.450	46.634	8.660	1.335	1.503
		Frozen green peas	USDA	1/2 cup	Number of 1/2 cup servings harvested for consumption[13]	14379.96	0.02					
		fresh green beans	USDA	1/2 cup	Number of 1/2 cup servings harvested for consumption[13]	129000	0.15					
		Yellow and white	USDA	1/2 cup	Number of 1/2 cup servings harvested for	323375.5	0.37					

		canned corn			consumption[13]							
Q19	Fruits	banana	USDA	1 medium piece	Number of servings consumed per capita[64-66]	64	0.13	0	65.258	17.090	0.187	0.613
		apple	USDA	1 medium piece		106.624	0.22					
		grapes	USDA	1/2 cup		300.672	0.63					
		canned fruit cocktail - light syrup	USDA	1/2 cup		7.0312	0.01					

BSQ Nutrition Database Analysis – Recipe defined by a single food

Line Item	Line Item Name	Recipe	Info. Source (USDA or other)	Serving Size	Consumption Information or Proxy		Weight	Added Sugars (g)	Energy (kcal)	CHO (g)	Fat (g)	Protein (g)
Q1	100% Fruit Juices	Apple Juice	USDA	1 cup	% of total fruit juice consumption[49]	35%	1.00	0	114.08	28.024	0.322	0.248
Q2	Fruit Drinks (e.g. Snapple)	Capri Sun	USDA	180 mL	\$1,000's spent on advertising, 2011[23]	12883	1.00	18.685	84.742	22.108	0.034	0.065
Q3	Sport drinks	PowerAde	USDA	20 fl oz	N/A	--	1.00	31.964	158.6	39.223	0	0
Q4	Flavored waters (e.g. Propel)	Propel	USDA	24 fl oz	N/A	--	1.00	0	0	0	0	0
Q5	Diet soda or pop	diet cola	USDA	13.16 fl oz	2011 market share of carbonated beverage brands, 85% soft drink consumption is 12 fl oz, 15% 20 fl oz, based on market shelf survey[20]	16.3	1.00	0.000	7.796	1.130	0.117	0.429
Q6	Regular soda or	cola	USDA	13.17 fl	2011 market share of carbonated	32.6	1.00	36.237	149.475	38.621	0.081	0.283

	pop			oz	beverage brands, 85% soft drink consumption is 12 fl oz, 15% 20 fl oz, based on market shelf survey[20]							
Q7	Energy drinks (e.g. Red Bull)	Red Bull	USDA	8 fl oz	2010 energy drink sales, \$1 millions[21, 22]	2,200	1.00	25.676	114.851	27.922	0.204	0.638
Q8	1% or nonfat milk	skim milk	USDA	8 fl oz	2006 milk sales, millions of lbs[50]	8,320	1.00	0.000	83.300	12.152	0.196	8.257
Q9	Regular or 2% milk	2% milk	USDA	8 fl oz	2006 milk sales, millions of lbs[50]	17,799	1.00	0.000	122.000	11.712	4.831	8.052
Q10	Low-fat or non-fat potato chips, tortilla chips and corn chips	Lowfat Potato Chips – salted	USDA	28g	2009 market share of potato chips (and related snacks) vs. tortilla chips, assume similar market shares among low-fat products[51-53]	64.1	1.00	1.873	115.398	23.001	1.998	1.850
Q11	Regular potato chips, tortilla	Regular potato chips - salted	USDA	28g	2009 sales, in \$1 millions[51-53]	7176.4	1.00	0.000	154.707	15.457	9.992	1.486

	chips, corn chips and puffs											
Q12	Other salty snacks (e.g. Chex Mix, Ritz Bitz)	Regular Butter Crackers	USDA	30g	2002 cracker sales, split evenly between cheese crackers, gold fish, butter crackers, \$1 millions[54]	1,118	1.00	0.000	152.750	18.226	7.623	2.666
Q13	Candy, including chocolate, candy bars, jelly bellies, gummies and Lifesavers	Snickers candy bar	USDA	58g	2009 sales, \$1 millions[56]	1720.2	1.00	28.556	277.998	35.920	14.017	4.002
Q14	Doughnuts, pop tarts or other breakfast pastries	regular pop tart	USDA	1 tart	2005 toaster pastry sales[57, 59]	480.00	1.00	12.550	211.140	37.384	5.702	2.522
Q15	Cookies, brownies, pies and cakes	Cake - Yellow, glazed, pudding type - added oil, unknown type of glaze, unknown egg,	USDA	80g (1 piece of cake)		2714	1.00	64.892	524.101	82.233	20.648	3.537

		unknown fat in frosting (prepared at home)										
Q16	Low or nonfat frozen desserts	Light ice cream - neapolitan, unknown fat %	USDA	1/2 cup (65g)	Amount of light ice cream / frozen yogurt / sherbert produced in 2009, 65g servings[62]	2,661	1.00	11.851	136.800	22.390	3.671	3.633
Q17	Regular ice cream and milkshakes	regular neapolitan ice cream	USDA	1/2 cup (65g)	Assume even distribution, unable to find milk shake consumption information[63]	0.34	1.00	8.850	138.598	16.587	7.260	2.376
Q18	Vegetables	Tossed Salad with dressing	USDA	1 cup	Number of 1 cup servings harvested for consumption[13]	395954.28	1.00	1.001	38.441	4.325	2.270	0.949
Q19	Fruits	grapes	USDA	1/2 cup	Number of servings consumed per capita[64-66]	300.672	1.00	0.000	52.095	13.666	0.121	0.544

BSQ Nutrition Database Analysis – Information provided by USDA and other sources

Line Item	Line Item Name	Recipe	Info. Source (USDA or other)	Serving Size	Consumption Information or Proxy	Weight	Added Sugars (g)	Energy (kcal)	CHO (g)	Fat (g)	Protein (g)	
Q1	100% Fruit Juices	Apple Juice	USDA	1 cup	% of total fruit juice consumption[49]	35%	0.38	0	126.960	30.788	0.316	0.925
		Orange Juice - calcium fortified	USDA	1 cup		31%	0.17					
		Orange Juice - vitamin fortified	USDA	1 cup			0.17					
		Grape Juice	USDA	1 cup		25%	0.27					
Q2	Fruit Drinks (e.g. Snapple)	Snapple	Other	16 fl oz	\$1,000's spent on advertising, 2011[23]	4,393	0.16	21.824	96.348	25.032	0.036	0.058
		Capri Sun	USDA	180 mL		12883	0.47					
		Kool Aid	Other	6.75 fl oz		10107	0.37					
Q3	Sport drinks	Gatorade	USDA	32 fl oz	N/A	--	0.50	21.824	96.348	25.032	0.036	0.058
		PowerAde	USDA	20 fl oz		--	0.50					
Q4	Flavored	Propel	USDA	24 fl oz	N/A	--	0.50	15.685	61.463	16.248	0	0

	waters (e.g. Propel)	Glaceau Vitamin Water Essential	Other	20 fl oz		--	0.50					
Q5	Diet soda or pop	diet cola with caffeine – medium	USDA	12 fl oz	2011 market share of carbonated beverage brands, 85% soft drink consumption is 12 fl oz, 15% 20 fl oz, based on market shelf survey[20]	16.3	0.76	0	6.938	1.006	0.105	0.426
		diet cola with caffeine – large	USDA	20 fl oz		16.3	0.13					
		diet non- cola – medium	USDA	12 fl oz		2.0	0.09					
		diet non- cola – large	USDA	20 fl oz		2.0	0.02					
Q6	Regular soda or pop	cola with caffeine – medium	USDA	12 fl oz	2011 market share of carbonated beverage brands, 85% soft drink consumption is 12 fl oz, 15% 20 fl oz, based on market shelf survey[20]	32.6	0.59	36.399	153.895	39.515	0.081	0.259
		cola with caffeine – large	USDA	20 fl oz		32.6	0.10					
		non-cola – medium	USDA	12 fl oz		14.3	0.26					

		non-cola – large	USDA	20 fl oz		14.3	0.05					
Q7	Energy drinks (e.g. Red Bull)	Rockstar	USDA	8 fl oz	2010 energy drink sales, \$1 millions[21, 22]	400	0.10	26.626	110.391	27.807	0.104	0.325
		Red Bull	USDA	8 fl oz		2,200	0.52					
		Monster	USDA	8 fl oz		1,500	0.36					
		Full Throttle	USDA	8 fl oz		110	0.03					
Q8	1% or nonfat milk	1% milk	USDA	8 fl oz	2006 milk sales, millions of lbs[50]	6,615	0.42	0.564	93.979	12.763	1.160	8.235
		skim milk	USDA	8 fl oz		8,320	0.53					
		1% milk - chocolate	USDA	8 fl oz		290	0.02					
		skim milk - chocolate	USDA	8 fl oz		365	0.02					
Q9	Regular or 2% milk	2% milk	USDA	8 fl oz	2006 milk sales, millions of lbs[50]	17,799	0.50	0.485	137.416	12.367	6.328	7.870
		whole milk	USDA	8 fl oz		16,410	0.46					
		2% milk - chocolate	USDA	8 fl oz		780	0.02					
		whole milk - chocolate	USDA	8 fl oz		719	0.02					
Q10	Low-fat or non-fat potato chips,	Lowfat Potato Chips – salted	USDA	28g	2009 market share of potato chips (and related snacks) vs. tortilla	64.1	0.80	1.498	116.250	22.718	2.173	1.868

	tortilla chips and corn chips	Lowfat-Baked Tortilla Chips – salted	USDA	28g	chips, assume similar market shares among low-fat products[51-53]	15.9	0.20					
Q11	Regular potato chips, tortilla chips, corn chips and puffs	Regular potato chips - salted	USDA	28g	2009 sales, in \$1 millions[51-53]	7176.4	0.48	0	149.938	16.349	8.826	1.560
		Regular tortilla chips - salted	USDA	28g		5579.8	0.38					
		Corn chips	USDA	28g		903.1	0.06					
		Cheese Puffs	USDA	28g		1217.4	0.08					
Q12	Other salty snacks (e.g. Chex Mix, Ritz Bitz)	Regular Cheese Crackers	USDA	30g	2002 cracker sales, split evenly between cheese crackers, gold fish, butter crackers, \$1 millions[54]	1,118	0.30	1.008	143.224	17.822	6.792	2.689
		Chex Mix – traditional	USDA	29g	2002 sales, "savory snacks - variety packs", \$1 millions[55]	346	0.09					
		Gold fish crackers - original	Other	30g	2002 cracker sales, split evenly between cheese crackers, gold fish, butter crackers, \$1 millions[54]	1,118	0.30					

		Regular Butter Crackers	USDA	30g	2002 cracker sales, split evenly between cheese crackers, gold fish, butter crackers, \$1 millions[54]	1,118	0.30					
Q13	Candy, including chocolate, candy bars, jelly bellies, gummies and Lifesavers	Snickers candy bar	USDA	58g	2009 sales, \$1 millions[56]	1720.2	0.43	15.500	218.412	28.992	10.967	3.012
		Hershey's Milk Chocolate Bar with Almonds	USDA	41g	2009 sales, \$1 millions[56]	1393.2	0.35					
		jelly beans - regular	Other	40g	2009 sales, \$1 millions[56, 67]	527.5	0.13					
		life savers	Other	15g	2009 sales, \$1 millions[56, 68]	362.922	0.09					
Q14	Doughnuts, pop tarts or other breakfast pastries	regular doughnut - cake, plain	USDA	55g	2005 doughnut sales, \$1 millions[57, 58]	242.67	0.20	9.861	222.769	30.727	9.729	3.158
		regular doughnut - raised, glazed, plain	Other	55g	2005 doughnut sales, \$1 millions[57, 58]	242.67	0.20					
		regular doughnut - raised, plain	USDA	55g	2005 doughnut sales, \$1 millions[57, 58]	242.67	0.20					

		regular pop tart	USDA	1 tart	2005 toaster pastry sales[57, 59]	480.00	0.40					
Q15	Cookies, brownies, pies and cakes	Cookies - chocolate chip, homemade from recipe, unknown fat	USDA	32g (2 medium cookies)	2010 sales of cookies / brownies / cakes / pies per store, \$1 millions [60, 61]	930	0.21	45.348	420.885	63.057	17.908	3.184
		Regular chocolate brownies, plain, not frosted, unknown fat - commercially prepared, unknown egg content	USDA	40g		138	0.03					
		Cake - Yellow, glazed, pudding type - added oil, unknown type of glaze, unknown	USDA	80g (1 piece of cake)		2714	0.61					

		egg, unknown fat in frosting (prepared at home)										
		Apple Pie - double crust, commercial	USDA	1 piece of pie (162g)		644	0.15					
Q16	Low or nonfat frozen desserts	Light ice cream - neapolitan, unknown fat %	USDA	1/2 cup (65g)	Amount of light ice cream / frozen yogurt / sherbert produced in 2009, 65g servings[62]	2,661	0.50	10.427	98.674	17.860	2.096	2.200
		Fat free sherbert	USDA	1/2 cup (85g)	Amount of light ice cream / frozen yogurt / sherbert produced in 2009, 85g servings[62]	282	0.05					
		Regular popsicle	Other	53g	Number of popsicles sold annually, millions of (1 popsicle) servings[48]	2,000	0.38					
		Low fat vanilla frozen yogurt	USDA	1/2 cup (100g)	Amount of light ice cream / frozen yogurt / sherbert produced in 2009, millions of 100g servings[62]	336	0.06					

Q17	Regular ice cream and milkshakes	regular neapolitan ice cream	USDA	1/2 cup (65g)	Assume even distribution, unable to find milk shake consumption information[63]	0.34	0.34	18.002	243.173	31.610	10.804	5.568
		milkshake - hard vanilla ice cream	USDA	1 cup (228g)		0.33	0.33					
		milkshake - soft vanilla ice cream	USDA	1 cup (186g)		0.33	0.33					
Q18	Vegetables	Tossed Salad with dressing	USDA	1 cup	Number of 1 cup servings harvested for consumption[13]	395954.28	0.46	0.450	46.634	8.660	1.335	1.503
		Frozen green peas	USDA	1/2 cup	Number of 1/2 cup servings harvested for consumption[13]	14379.96	0.02					
		fresh green beans	USDA	1/2 cup	Number of 1/2 cup servings harvested for consumption[13]	129000	0.15					
		Yellow and white canned corn	USDA	1/2 cup	Number of 1/2 cup servings harvested for consumption[13]	323375.5	0.37					
Q19	Fruits	banana	USDA	1 medium piece	Number of servings consumed per capita[64-66]	64	0.13	0	65.258	17.090	0.187	0.613

	apple	USDA	1 mediu m piece	106.624	0.22				
	grapes	USDA	1/2 cup	300.672	0.63				
	canned fruit cocktail - light syrup	USDA	1/2 cup	7.0312	0.01				