Published in final edited form as:

J Food Compost Anal. 2015 November 1; 43: 7–17. doi:10.1016/j.jfca.2015.04.004.

Estimating added sugars in US consumer packaged goods: An application to beverages in 2007–08

Shu Wen Ng, PhD^{*,a}, Gregory Bricker, MSc^a, Kuo-ping Li, PhD^a, Emily Ford Yoon, MPH RD^a, Jiyoung Kang, MSc RD^{a,b}, and Brian Westrich, PhD^c

^aUniversity of North Carolina at Chapel Hill CB #8120, Chapel Hill, NC 27599

^bJK conducted this work while with UNC-Chapel Hill

^cMcWest Corporation 2321 30th Ave S, Minneapolis, MN 55406

Abstract

This study developed a method to estimate added sugar content in consumer packaged goods (CPG) that can keep pace with the dynamic food system. A team including registered dietitians, a food scientist and programmers developed a batch-mode ingredient matching and linear programming (LP) approach to estimate the amount of each ingredient needed in a given product to produce a nutrient profile similar to that reported on its nutrition facts label (NFL). Added sugar content was estimated for 7021 products available in 2007–08 that contain sugar from ten beverage categories. Of these, flavored waters had the lowest added sugar amounts (4.3g/100g), while sweetened dairy and dairy alternative beverages had the smallest percentage of added sugars (65.6% of Total Sugars; 33.8% of Calories). Estimation validity was determined by comparing LP estimated values to NFL values, as well as in a small validation study. LP estimates appeared reasonable compared to NFL values for calories, carbohydrates and total sugars, and performed well in the validation test; however, further work is needed to obtain more definitive conclusions on the accuracy of added sugar estimates in CPGs. As nutrition labeling regulations evolve, this approach can be adapted to test for potential product-specific, category-level, and population-level implications.

Keywords

Food composition; Food analysis; Added sugars; Linear programming; Nutrition label; Ingredients; Beverages

***Corresponding Author:** Shu Wen Ng, Department of Nutrition, Gillings School of Global Public Health, University of North Carolina at Chapel Hill. Mail: CB # 8120 University Square, Chapel Hill, NC 27516. Tel.: (919) 962-6188. shuwen@unc.edu. gbricker@email.unc.edu, kpli@email.unc.edu, ebford@email.unc.edu, jiyoung.kang3@gmail.com, bw@mcwest.com

Publisher's Disclaimer: This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final citable form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

Conflicts of Interest

Westrich has in the past and may in the future provide consulting services related to the use and development of linear programming approaches for estimating nutrient values in food composition databases to the US Department of Agriculture and other organizations. Otherwise none of remaining the authors have any conflicts of interest to declare.

1 Introduction

Added sugars, that is sugars included in foods during processing or preparation, comprise the majority of sugars in the typical American diet (R. K. Johnson et al., 2009; Reedy, 2010; U.S. Department of Agriculture, 2010; U.S. Department of Agriculture & U.S. Department of Health and Human Services, 2010). While chemically indistinguishable from naturally occurring sugars (e.g. fructose in fruit or lactose in milk and milk products), added sugars have become an ingredient of public health concern. Foods containing high amounts of added sugars are often sources of energy with very few nutrients (e.g. sugar-sweetened beverages, grain-based desserts, dairy desserts, and candy) (Fitch & Keim, 2012; Ng, Slining, & Popkin, 2012). Overconsumption of these foods may lead to excess energy intake and poor diet quality (U.S. Department of Agriculture, 2010). Moreover, when considered with solid fats and excess energy intake, added sugars can potentially lead to adverse health effects, including obesity, type-2 diabetes or pre-diabetes, inflammation, and cardiovascular disease (Rachel K. Johnson et al., 2009; Malik et al., 2010; Morenga, Howatson, Jones, & Mann, 2014; Te Morenga, Mallard, & Mann, 2013; Welsh, Sharma, Cunningham, & Vos, 2011).

Appropriately, the Dietary Guidelines for Americans (DGA) include recommendations to reduce the intake of calories from added sugars (U.S. Department of Agriculture, 2010; U.S. Department of Agriculture & U.S. Department of Health and Human Services, 2010). However, it is difficult for consumers to adhere to these recommendations, as the amount of added sugars cannot be identified from the nutrition facts label (NFL) on consumer packaged goods (CPGs). While there is growing concern over the use of added sugars in the US food system, monitoring their presence in products and their consumption remains challenging for the following reasons: a) no laboratory method can analyze for added sugars (not chemically distinguishable from naturally occurring sugars); b) current nutrition labeling regulations do not require that added sugars be reported separately from total sugars; c) added sugar amounts must be either estimated or supplied by food companies; and d) estimations must be able to keep up with new and reformulated products in order to capture changes in the food system.

In February 2014, the Food and Drug Administration (FDA) released a proposed update to the nutrition labeling regulations that includes a mandatory disclosure of added sugar content to be listed on the NFL (Food and Drug Administration, 2014b). In addition, it sought to define added sugars as: “sugars that are either added during the processing of foods, or are packaged as such, and include sugars (free, mono- and disaccharides), syrups, naturally occurring sugars that are isolated from a whole food and concentrated so that sugar is the primary component (e.g. fruit juice concentrates), and other caloric sweeteners... Sugar alcohols are not considered to be added sugars...” (Food and Drug Administration, 2014b).

The proposed revision to the nutrition labels is not finalized, so it is unclear what changes will be made and when any revisions to the nutrition labels would come into effect.

We designed a batch-mode approach to estimate added sugar content in commercially formulated CPG products using linear programming (LP), with an application to CPG beverages available in 2007–08. This followed work undertaken by the University of Minnesota Nutrition Coordinating Center and the United States Department of Agriculture (USDA), both of which have used LP approaches to estimate missing nutrient values (Schakel, Buzzard, & Gebhardt, 1997; Westrich, Altmann, & Potthoff, 1998; Westrich, Buzzard, Gatewood, & McGovern, 1994) for food composition tables. We built off these past LP methods, and estimated added sugars content in products using a systematic batch-mode manner that allows for larger-scale applications. Because there is no cost- or time-efficient gold-standard in which to compare our results across thousands of products, we assessed the validity of our estimated nutrients in two ways: a) by comparing LP estimated nutrient values to known values from the NFL, and; b) by conducting a small validation study based on 15 known formulations.

2 Materials and methods

2.1 Overview

The process for estimating added sugar values requires three pieces of information: the nutrition facts label (NFL), the ingredient list, and nutrient composition for each ingredient. We utilized a linear programming (LP) approach to estimate the amount of each ingredient needed in a given product to produce a nutrient profile as close as possible to that reported on its NFL. To help develop accurate estimations, constraints were applied to ingredients using information gathered from FDA labeling laws, scientific journals, and knowledge of typical manufacturing processes. Once formulations were estimated, we calculated the added sugars content of each product by summing the amount of sugar from ingredients identified as added sugars.

2.2 Data Sources

2.2.1 Nutrition facts label (NFL) Label and Ingredient Statements (UPC level data)—The 2007 and 2008 NFL data came from a number of sources, described in detail in other papers (Ng & Popkin, 2012; Ng, et al., 2012). These data provide nutrition information for uniquely formulated commercial foods and beverages at the Universal Product Code (UPC) level. Per FDA requirements, most NFL labels must include the serving measurement, total calories, calories from fat, total fat, saturated fat, trans fats, total sugars, total carbohydrate, protein, dietary fiber, sodium, cholesterol, vitamin A, vitamin C, calcium, and iron (National Archives and Records Administration, 1993).

2.2.2 Ingredient Nutrient Profiles (Ingredient level data)—We used the 2007 version of ESHA Ingredient database (ESHA Research), which is a comprehensive and key ingredient nutrition profile database used in the food-manufacturing industry. This database includes >37,000 foods and food items with data from >2600 sources. Data sources include USDA Standard Reference Database, manufacturer's data, restaurant data, and data from scientific literature. The database also provides nutrient information for commonly used industrial ingredients (including gums, preservatives, and vitamins and minerals). In addition, we created additional ingredient nutrient profiles for ingredients not in the ESHA

database by referring to ingredient supplier websites and food science reference books (BASF, online; Farbest Brands, online; Food and Agriculture Organization, 2006a, 2006b; Ockerman, 1991). Lastly, there are often multiple nutrient profiles in the ESHA database for a single ingredient. In these cases, the nutrient profiles were averaged and used as the nutrient profile for the ingredient.

2.3 Data Preparation

A team with food science, dietetics, biostatistics and statistical programming backgrounds collaborated to clean, manage and process the data from both the NFL and ingredient level databases. Preparation steps for the data sources included: ingredient list parsing; converting NFL and ingredient nutrient profiles to reflect nutrients per 100 grams and; identifying added sugars within the ingredient level database.

2.3.1 Ingredient list parsing—FDA regulations require that ingredients be listed in order of weight if they contribute more than 2% of the formulation. We programmatically parsed out ingredients and ‘sub-ingredients’. For example, at times, ingredients might be listed as “fruit juice blend from concentrate (water, apple juice concentrate, pear juice concentrate)¹” or “stabilizer (carrageenan)”. In these cases, the hierarchical order of the ingredient and sub-ingredients were maintained as the additional information can be useful for the ingredient linking process². Additionally, this parsing process identified when listed ingredients follow statements such as “contains 2 (or 1.5 or 1) percent or less of”. In these cases, we flagged these ingredients as ‘manufacturer 2% ingredients’ for each product.

2.3.2 Converting NFL and ingredient nutrient profiles into per 100 grams—The NFL data is typically reported in terms of per serving. Along with information about the total weight or volume of the product, we converted the NFL data into per 100g, after accounting for density factors when necessary. Similarly, we converted the ingredient nutrient profile data into per 100g. This ensured that all estimations were conducted within the same unit of measure.

2.3.3 Identifying added sugars within the ingredient database—Following the FDA proposal, all free, mono- and disaccharides and syrups were considered added sugars (including: agave syrup, brown sugar, corn syrup, corn syrup solids, dextrose (glucose), fructose, high fructose corn syrup, honey (liquid and dry), invert sugar, lactose, malt syrup, maltose, maple syrup, molasses, rice syrup, sucrose, trehalose, turbinado sugar), but sugar alcohols were not. While fruit juice concentrates (FJC) have been identified as sources of added sugars, it was unclear how to appropriately distinguish products that contain fruit juices from concentrate from those that contain both FJC and water. Even in the recent proposed FDA regulations that considers “naturally occurring sugars that are isolated from a whole food and concentrated so that sugar is the primary component (e.g. fruit juice

¹The FDA food labeling guide on juices states that is incorrect to group a blend of juice concentrates (Food and Drug Administration, 2013), so this example is a non-compliant ingredient list. However, these do exist and need to be handled, so we applied the rule we described.

²In the first example, because it is unknown what proportion of the fruit juice blend from concentrate came from apple vs. pear, we use the first listed juice (i.e. we used the ingredient nutrient profile for apple juice from concentrate).

concentrates)” as added sugars, it is not clear what constitutes “primary component” and how to categorize juices from concentrates. In this study on beverages, we did not consider fruit juices from concentrate or FJC as added sugars. However, we tested how results from a subset of beverages might vary if FJC was included as an added sugar and if both FJC and juice from concentrate were included as sources of added sugars (see Section 2.6).

2.4 Batch-Mode Linear-Programming (LP) Approach for Estimating Added Sugar Values

All commercial foods and beverages within the product database are identified by a UPC, and contain package label information, including brand name, ingredient list, NFL, and manufacturer claims. The food scientist and registered dietitians reviewed the product attributes for beverage products and identified ten distinct categories, or batches, that encompassed all products to be run through the LP. In a two-part process, each ingredient occurring in a batch was linked to an ingredient nutrient profile. The ten batches of products were run through a linear programming (LP) model to estimate food formulations and thus added sugar values. An overview of the approach is summarized in Figure 1. All data preparation was conducted in Excel (Microsoft, 2013) and SAS version 9.3 (SAS Institute Inc., 2011). The linear programming model was conducted in SAS/IML version 12.1 (SAS Institute Inc., 2012).

2.4.1 Identifying batches of products with similar forms and manufacturing practice—For ease of data management and handling, products were classified into broad, aggregate commercial categories that could be disaggregated into batches based on their nutritional likeness and their location within the grocery store (e.g. ready to drink, liquid concentrate, powder concentrate). Once a batch was identified, a food scientist or registered dietitian conducted additional reviews of product attributes to ensure they had similar form and manufacturing practices. This helped ensure the generalizability of ingredient nutrient profiles across multiple products in a batch given the nature of the batch-mode linking process and FDA regulations regarding ingredient designation. We excluded batches of beverages that did not contain sweeteners (e.g. plain dairy milks, unflavored waters), and single ingredient products.

2.4.2 Matching the Most Appropriate Ingredient Nutrient Profiles to NFL Ingredients—Before a NFL listed ingredient was matched with an ingredient nutrient profile, extensive research of FDA labeling regulations, manufacturing practices, and food science literature was performed for each batch of interest. This information provided guidance for matching nutrient profiles to ingredients listed on the NFL, and served as the basis for many of the constraints applied to the LP model (described below). Ingredient nutrient profiles were selected based on the final form of the product. For example, if ‘milk’ was listed on the ingredient list of a powdered dairy drink mix, it was matched to an ingredient nutrient profile for powdered milk, not liquid milk.

The process of matching ingredient nutrient profiles to each NFL ingredient can be time-consuming when handling each UPC separately. In order to expedite this process, we developed a two-part approach to programmatically assign nutrient profiles to ingredients. Briefly, after a batch is identified, two files of parsed ingredients are created. File A contains

a list of unique ingredients sorted by their frequency of occurrence within the batch, while File B contains parsed hierarchical ingredient (and sub-ingredient) lists and product attributes for each UPC within the batch. The primary objectives of File A are to:

- match each uniquely occurring ingredient to an appropriate ingredient nutrient profile code to minimize linking common ingredients repeatedly and ensure consistency;
- apply constraints to specific ingredients that are always 2% of a product formula even when manufacturers do not identify them as such (e.g. ingredients with maximum legal usage rates like certain preservative, vitamins and minerals; or ingredients that are always used at low levels to provide functional benefits like gums and other stabilizers)—we refer to these as ‘researcher 2% ingredients’, and;
- rename synonymous ingredients when appropriate.

Nutrient profile codes, constraints or renames applied to ingredients in File A are then used to populate File B ingredients. After a review and product-specific adjustments by a registered dietitian or food scientist, the finalized File B is used as one of the input files for the LP model.

2.4.3 Linear Programming (LP) Model Estimations—Conventional methods used to estimate unknown nutrient values in commercially formulated foods and beverages require the ingredient information (listed in descending order of weight), nutrient profiles for all ingredients listed, nutrient information for the food mixture, and any other known constraints. When this information is known, product formulations can be estimated by adjusting the proportions of ingredients until the sum of nutrients matches the known nutrients of the composite food (Schakel, et al., 1997; Westrich, et al., 1998; Westrich, et al., 1994).

For each batch, we were able to conduct LP estimation because we had the necessary information including:

- NFL ingredients listed in descending order, with flags on which ingredients are ‘manufacturer 2%’ or ‘researcher 2%’, for each UPC along with the matched ingredient nutrient profile code for each ingredient [final File B].
- Ingredient nutrient profile database with ingredients identified as added sugars or not.
- UPC-level NFL key nutrient information per 100g on calories, total fat, saturated fat, cholesterol, protein, carbohydrate, total sugar, fiber, sodium and potassium [‘known’ nutrient and weight (100g) values].

The objective function (goal) of the LP model is to minimize the total model ‘error’, subject to a number of constraints. The choice of how errors are expressed is not straightforward. Expressing errors as percentages (in relative terms) will overemphasize differences for nutrients that are present in products in small quantities. For example, consider a milk drink that contains 2g fat, 10g carbohydrates and 250mg sodium per 100g, a 1g difference in fat is considered a 50% error, which would be treated with equal significance as a 5g difference in

carbohydrates. Using absolute differences on the other hand also has its drawbacks. In the above example, the carbohydrate difference of 5g would be considered 5 times more significant than the fat difference of 1g. In addition, nutrient units of measure can change the relative emphasis placed on different nutrient errors. For example, the numeric value of sodium is often high because it is measured in milligrams, while the numeric value of fat is typically lower as it is measured in grams. So in the above example, a 50mg difference in sodium would be considered to be 50 times more important than the fat error.

Consequently, we used nutrient error tolerances in order to allow comparison of biological significance of nutrient estimation errors across different nutrients. We based these nutrient error tolerances values on those used in past studies that also use an LP modeling approach for estimating missing nutrient values (Westrich, et al., 1994), which are typically based on 5-10% of recommended Daily Values (DV). This allows us to scale different nutrients with different units of measure and biological significance into comparable tolerance units.

We also included a weight error tolerance of $\pm 5\text{g}$ as part of the LP objective function in order to handle ingredients that do not contribute any nutrients but do contribute weight, such as water. The choice of a $\pm 5\text{g}$ weight error as being equivalent to 1 tolerance unit was used as a starting point as the National Institute of Standards and Technology (2005) suggests that the maximum allowable variance should be package weight specific. We do not mean to imply that the biological significance of 5g (of any product) is similar to that of 85 kcals (the nutrient error tolerance used for calories), but since the nutrient tolerances were based on 5-10% of recommended DVs, we chose a conservative value as the weight tolerance. Future improvements to the LP model might include taking into account product specific weight error or maximum allowable variance as being equivalent to 1 tolerance unit. Meanwhile, Table 1 presents the weight and nutrient errors applied in this paper to be equivalent to 1 tolerance unit.

Therefore, the objective function of the LP model is to minimize the sum of the absolute values of the weight and nutrient errors as described in Equation (1):

$$\text{minimize } Z = \sum_k (p_k + q_k), \quad (1)$$

where Z is the total model error, p_k and q_k are the negative (under estimation) and positive (overestimation) errors respectively for the k th weight or nutrient, expressed in tolerance units.

Equation (1) is subject to the following constraints:

- ingredients are listed on the NFL in descending order of predominance by weight (order constraint)³;
- ingredients listed as 2% by manufacturers are constrained to account for 2% of total weight of the product, and ingredients listed before the first 'manufacturer 2%' ingredient are constrained to account for >2% of total weight of the product. ('manufacturer 2%' constraint);

³The order constraint is removed for 'manufacturer 2%' ingredients, but is retained for 'researcher 2%' ingredients.

- ingredients not listed on the NFL as 2%, but used at less than 2% are also constrained to only account for 2% of total weight of the product ('researcher 2%' constraint);
- every ingredient contributes non-negative weight to the product (0% constraint).

As shown in Figure 1, the LP model used the parameters and input data described above to estimate the amount of each ingredient present in 100g of a product. Based on the estimated formulation, we calculated the added sugar in grams and in calories per 100g of the product using total sugar grams or calories from ingredients identified as added sugars. Total weight of the ingredients identified as added sugars was not used in the calculations because not all of the weight or calories from added sugar ingredients are sugar. For example, honey is an added sugar that contains 82g sugar per 100g. Thus, if 5g of honey are used per 100g of product, the added sugar contribution from honey is 4.1g per 100g product.

To provide an illustration, Figure 2 explains how the NFL data, ingredient nutrient profile matching and LP model come together to estimate the formulation for a very simple peanut butter product. This is a simplified example since the only source of protein is peanuts in the ingredient list. Most products are much more complex with regards to the various sources of each nutrient. This is why the LP approach which takes into account the various constraints can achieve great efficiencies in estimating formulations. Additionally, Appendix A (online Supplementary data) illustrates how Z (total model error) is calculated for a more complex Chocolate Milk example after its formulation is estimated.

We present added sugar results for ten beverage categories: caloric sodas & energy drinks; sports drinks; flavored waters-carbonated & still; fruit & vegetable juice drinks (beverages that contain any amount of fruit/vegetable juice); fruit flavored drinks (does not contain fruit juice, only fruit flavor); sweetened dairy & dairy alternatives; ready-to-drink (RTD) coffees & teas; concentrated fruit drinks; powdered beverage mixes, and; cocktail mixes. We report the number of products included in each of these key beverage categories of interest, the LP estimates compared to the NFL nutrient and weight values in both absolute and relative terms⁴, and the added sugar estimates per 100g. We also provide statistics on the median total error per nutrient (standardizes the total model error Z by the number of known nutrients and weight in each product).

2.5 Validation of LP estimations

Short of obtaining official formulations of products from 2007–08 from manufacturers, it is not possible to properly validate the estimated formulations and added sugar values derived from the LP model. Therefore as a small validation test, the food scientist on our team created 15 product formulations of six types of beverages that contained both intrinsic and added sugars. These formulations are meant to reflect items available in the marketplace. Then, an independent NFL labeling firm generated NFL data for these 15 products, and we ran these products as a 'validation batch' through the LP model and obtained estimated formulations and nutrient values per 100g. The use of an independent NFL labeling firm

⁴For products with missing or zero nutrient values on the NFL, we are unable to conduct relative comparisons of the LP estimates to the NFL values.

meant that the ingredient nutrient profile database used by them and the decision of which nutrient profile to link to each ingredient was external to us and avoided circular analysis. We then compared how these estimated values compared to the known formulations as well as the NFL nutrient values provided by the independent labeling firm.

2.6 Sensitivity analyses

The estimated formulations and thus added sugar values may vary depending on the parameters used in the LP model, so we conducted a number of sensitivity analyses to ensure that the findings were robust. First, we tested if distinguishing between ‘manufacturer 2%’ and ‘researcher 2%’ ingredients in the constraints would affect the added sugar estimates. We repeated these within the small validation test of the 15 formulations as well, and present these results. Second, as mentioned previously, it is unclear whether FJC and juice from concentrate should be considered sources of added sugar. Therefore, we calculated estimated added sugar values under various definitions for five beverage categories of concern: caloric sodas & energy drinks, flavored waters, fruit & vegetable juice drinks, fruit flavored drinks, and sweetened dairy & dairy alternatives. This can show how different definitions regarding what is considered an added sugar would affect the estimated proportion of products that contain added sugars and estimates of the amount of added sugars in the food supply.

3 Results

There were 7021 products from ten beverage categories in the product database from 2007–08 that were candidates for this analysis (total sugar values >0g per 100g): caloric sodas & energy drinks (n=1711), sports drinks (n=290), flavored waters-carbonated & still (n=331), fruit & vegetable juice drinks (n=1649), fruit flavored drinks (n=183), sweetened dairy & dairy alternatives (n=860), RTD coffees & teas (n=742), concentrated fruit drinks (n=130), powdered beverage mixes (n=942) and cocktail mixes (n=183). Table 2 presents the select results for each of the ten beverage categories, while Appendix B (online Supplementary data) presents the detailed results.

We were able to estimate non-zero added sugar values for almost all of the caloric sodas and energy drinks, which resulted in an average of 10.8g of added sugars per 100g of product (sd=2.31) that contributed 98% of the calories in these products. Across the ten beverage categories, flavored waters had the lowest absolute amounts of added sugars (mean= 4.3g/100g), although there was a large variance in the added sugar content of these products (sd=3.2). Even with a low mean absolute amount of added sugar, they still contributed to 92% of the caloric content of these products. We found similar results for sports drinks (mean added sugar = 5.2g/100g, which contributed close to all of the calories).

Among fruit and vegetable juice drinks, when FJCs are not defined as added sugars, mean added sugar content was estimated to be 8g/100g (sd=3.84), and contributed 71% of calories. Meanwhile, fruit flavored drinks had similar added sugar content (mean=7.7g/100g), which contributed 97.5% of calories.

Our estimates show that sweetened dairy & dairy alternatives contain 6.2g of added sugars per 100g, but because of the calories from protein and fat in these products, added sugars made up 33.8% of calories, which was the lowest percentage of calories from added sugars among the categories studied. Estimated added sugars in RTD coffees and teas constituted 7.2g/100g and accounted for 79.6% of calories.

Figure 3 shows the gram contribution of intrinsic vs. added sugars to total sugars across RTD beverage categories among UPCs with NFL total sugars >0g when fruit juice and FJCs are not considered added sugars. Flavored fruit drinks had the largest percentage of total sugars from added sugars (100%), while sweetened dairy & dairy alternatives had the smallest percentage of added sugars (65.6% of total sugars).

The three non-RTD beverage categories (concentrated fruit drinks, powder beverage mixes and cocktail mixes) had considerably higher content of added sugars/100g (36.1g, 66.2g, and 19.3g, respectively) since they are not yet reconstituted for consumption. At the same time, added sugars contributed a lower share to the caloric content of these products (77.6%, 67.7%, and 84.4%, respectively) compared to the other beverage categories, likely because they also include ingredients such as FJC and milk powders.

As shown in Table 2, LP estimates were comparable to NFL values for calories (mean LP-to-NFL ratio: 0.99-1.08), carbohydrates (mean ratio: 1-1.08), and total sugars (mean ratio: 0.93-1.04). However, the results for total fat, saturated fat and protein for most of the beverage categories had more variable LP-to-NFL ratios due to the smaller number of products having non-zero or non-missing values for these nutrients.

The median total errors per nutrient were generally low for RTD beverages, but were higher for non-RTD beverages, particularly concentrated fruit drinks (0.065 tolerance unit) and powdered beverage mixes (0.136 tolerance unit). This is expected since the serving sizes for such products are smaller and so FDA nutrition label rounding rules result in greater imprecision.

3.1 Validation results

Appendix C (online Supplementary data) shows the product-level results from the small validation test on 15 product formulations across six beverages. Select product formulations are presented in Table 3. These results show that the LP model distinguished between varying amounts of ingredients in the formulations. However, the LP model performed less well for products that contain multiple similar ingredients that are listed in close proximity to each other, such as in the case of the vegetable juice drink where the tomato juice from concentrate and the carrot juice for concentrate are listed sequentially⁵. For the 15 products, the estimated LP to actual added sugar differences ranged from 0 to 1.77g per 100g of the product formulation.

⁵In general, the LP model only differentiates between ingredients based on their nutrient profile and their order in the ingredient list. When adjacent ingredients have a similar nutrient profile, there is no distinguishing information that can be used by the LP model to determine the relative proportion of these ingredients to each other (other than the order constraint). Thus, the estimated relative proportion of these ingredients should be considered to be fairly arbitrary.

We also see from Appendix C that the estimated results did not appear to change meaningfully whether the ‘manufacturer 2%’ constraint is applied or not. There was also no obvious bias between the two; in some cases the values were identical, but there were also cases where including the ‘manufacturer 2%’ constraint produced higher added sugar estimates compared to excluding it, as well as vice versa.

3.2 Fruit juice concentrates and juice from concentrate

Table 4 presents the results for the number of products that contain any added sugars, their content, and contribution to total sugars under three definitions of added sugars. As expected, these all rose as the definition widened to include FJCs and then also FJC and fruit juices. Not surprisingly, the largest difference on the number of products containing any added sugars and their added sugar content was for fruit & vegetable juice drinks as the definition expanded, and the least difference was for sweetened dairy & dairy alternatives.

4 Discussion

This batch-mode approach to estimate added sugars content in consumer packaged goods (CPG) products delivered results that appear reasonable and robust. When applied to 7021 qualifying CPG beverages products in the 2007–08 product database, 6729 (95.8%) contained added sugars. Added sugar accounted for 65.6% to 100% of total sugars across the ten beverage categories studied. Additionally, the LP estimates for total sugar, carbohydrates and calories compared well to the NFL values across these beverage categories, while the LP estimates for total fat, saturated fat and protein performed reasonably for the sweetened dairy and dairy alternative category. A small validation study supported our findings that this method can be used to estimate added sugar levels in beverages well. We also show how altering the added sugar definition will affect results.

Recently, the FDA proposed updates to the NFL label requirements, including the addition of added sugars (Food and Drug Administration, 2014a). This proposal is currently under revision based on public comments, so there is no guarantee that added sugars will be included in the final regulations, and it will likely be several years before any update will come into effect. Regardless, there is no historical information about the added sugars of products and therefore, no way to determine if changes in the formulation of CPG foods and beverages are occurring or moving in the desired direction.

Additionally, the approach described here provides some flexibility in adjusting parameters in the LP model as FDA regulations evolve. For example, should regulations define an ingredient previously considered an added sugar to no longer be considered one (or vice versa), we can change this in the relevant input file and generate new results, as illustrated in Table 4. Similarly, we could test proposed definitions for added sugars to understand the breadth and degree of their effects.

In the FDA proposal, FJC is included as an added sugar because it is isolated from a whole food and concentrated. Prior studies that examined ingredient lists in NFL databases found that 7% of 85,451 unique CPG food and beverages contained FJC, and 75% of those CPG products contained some type of added sugars (Ng, et al., 2012). The added sugar estimation

approach described here goes further by allowing us to not only determine the proportion of CPG products that contain added sugars, but also estimate the amount present. This is the first effort to begin documenting the degree of added sugars in the US CPG food supply. Future work will also combine existing efforts to link UPCs to USDA food codes (Slining, 2015) to estimate the presence and amount of added sugar in both purchases made by US households as well as in foods consumed.

Currently, we have completed the work for CPG beverages. We have begun expanding this work into CPG foods to establish batch-specific ingredient matches and include potential changes to the LP parameters as necessary. We plan to repeat the established protocols with new CPG products and formulations available in future years. If and when the new proposed labeling regulations by the FDA come into effect, we may compare the results of analyses conducted using data from similar time periods to test and enhance this approach.

It is also possible to expand this approach to estimate other nutrients. However, adjustments to the LP parameters will be necessary based on FDA labeling regulations on those nutrients and data availability for both NFL and ingredient nutrient profile databases.

Lastly, in order to be careful to not overstate our findings, we chose to err on the side of minimizing added sugar estimates in all decisions made during data preparation, and in the applying of tolerances and constraints. Thus the results presented here should be considered minimum estimated added sugar amounts. Future work could take the opposite approach to estimate maximum added sugar amounts. This would provide us with lower and upper bounds for added sugar estimates and determine how large the range is.

4.1 Limitations

This work relies on commercial databases and therefore is a costly endeavor. However, in comparison to the costs of laboratory analyses, this approach can become cost efficient once protocols are established for various batches of CPG products because of the number of products that the LP model can be applied to. Regardless, in the case of added sugars, laboratory analyses would be unable to distinguish naturally occurring and added sugars.

We found that we were either unable to estimate the added sugar content for products or had problematic estimates when there were substantial inaccurate or incomplete data in the input files (NFL nutrients, NFL ingredients or ingredient nutrient profiles). Unfortunately, it is not simple to determine to what extent these occurred. However, in the data preparation process, we conducted quality and integrity checks on the NFL nutrient and ingredient data as best possible to optimize their accuracy following earlier examples (Murphy, 1989). In addition, we are able to impute nutrient values in some cases (e.g. when there was data on calories, fat and carbohydrate, we derived missing protein values). With regards to the ingredient nutrient profile data, although the nutrient profiles are reviewed, we did not impute any missing values.

Moreover, even if the NFL databases do accurately reflect labels, it is possible that the labels may not properly reflect the nutrient content of the products. Past studies have found some large discrepancies between values on NFLs and what was found during laboratory analyses

(Food and Drug Administration, 2003; Urban et al., 2010). Our study is only able to provide insight on the added sugar content based on labels and assumes that on average, the labels do adhere to FDA regulations and are within the bounds of allowances.

The NFL nutrient data for products with small serving sizes (e.g. powder coffee creamers) also posed significant challenges. We are unable to estimate the formulations and added sugar content of many of these products because of rounding allowances in labeling regulations. With small serving sizes, many nutrients will be very close to zero and often are reported as such, resulting in systematic inaccuracies in known nutrient values⁶. As shown in the estimates for powdered and concentrate beverage products, these estimates did not perform as well compared to the RTD beverage products.

A key constraint in the LP model is the ordering of ingredients listed on labels. Consequently, we are unable to estimate added sugar content for: products whose ingredient list included 'super-ingredients' (e.g. a frozen cake with separate ingredient lists for the icing and the cake), or products with complex, multi-component ingredients that cannot be linked to a single ingredient nutrient profile (e.g. Strawberry Base (Sugar, Strawberries, Water, Natural Flavor, Annatto [Color], Red 40, Citric Acid)). We also cannot estimate added sugars for variety packs because there is more than one NFL label and ingredient list per UPC. However, items within variety packs are likely to also be sold separately and thus likely to be captured that way.

Since the focus here is on added sugars, decisions made along the way were aimed at ensuring that estimates for added sugars and its associated nutrients (i.e. total sugars, carbohydrates) were appropriate. In some cases this may have resulted in poorer estimates for other nutrients such as protein or fat, particularly for those with unreported NFL values for these nutrients.

Despite these limitations, we were still able to identify and apply the LP model for over 7000 CPG beverage products sold in 2007-08. The systematic approach of mapping ingredients and conducting LP estimations in a batch-mode minimizes variability in the decisions made, and allows for replicability.

5 Conclusions

In conclusion, our findings suggest that using linear programming to estimate added sugar contents of products in a systematic batch-mode manner is possible and has a number of advantages. This approach allows for flexibility in setting and testing the parameters applied across multiple products at once. For example, should regulations regarding the definition of added sugars change, it would be possible to adapt the identification of added sugar ingredients and allow a re-estimation of added sugar amounts in products and the food supply, which we demonstrate in a small sample. This method will also allow for more timely measurements of the use of added sugars in the US food supply as well as its

⁶Rounding errors can also be introduced when converting to per 100g amounts if the serving size <100g. For example, a reported 4g sugar per 20g serving is converted to 20g sugar/100g product. However, due to rounding, the actual sugar amount could be 3.6g sugar/20g product = 18g sugar/100g product or 4.4g sugar/20g product = 22g sugar/100g product.

prevalence in CPG products purchased and consumed. It also has the potential to enhance the capability of food composition databases to continually review and update estimated nutrients, such as added sugars, particularly for CPG foods and beverages. Expanding this approach to other time periods, nutrients and countries will be less costly once protocols for ingredient matching and LP parameterization are established. However, sufficient high quality NFL and ingredient nutrient profile data is critical before this approach becomes scalable.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

We thank the National Institutes of Health R01DK098072 and CPC (R24 HD050924) for financial support. We also wish to thank Barry M. Popkin for his help in conceptualizing this project and comments on the paper, and Meghan Slining, Jessica Davis, Bridget Hollingsworth, Julie Wandell and Jim Terry for assistance in this effort. We dedicate this work to the memory of Dan Blanchette. Other than noted above, none of the authors have conflict of interests of any type with respect to this manuscript.

Abbreviations

CPG	Consumer packaged goods
DV	Daily values
DGA	Dietary Guideline for Americans
FDA	Food and Drug Administration
NFL	Nutrition facts label
RD	Registered Dietitians
RTD	Ready-to-drink
UPC	Universal product code
USDA	United States Department of Agriculture

References

- BASF. Newtrition: Food, Beverage and Dietary Supplements. online http://www.newtrition.basf.com/web/global/newtrition/en_GB/market_segments/index
- ESHA Research. Esha Ingredient Database, 2011. 2008. from <http://www.esha.com/solutions/additional-databases/>
- Farbest Brands. Ingredients. online <http://www.farbest.com/ingredients.asp>
- Fitch C, Keim KS. Position of the Academy of Nutrition and Dietetics: Use of Nutritive and Nonnutritive Sweeteners. Journal of the Academy of Nutrition and Dietetics. 2012; 112(5):739–758. [PubMed: 22709780]
- Food and Agriculture Organization. Combined Compendium of Food Additive Specifications. 2006a. (online database and report) (Publication no. ISBN 92-5-105569-6). Retrieved <http://www.fao.org/docrep/009/a0691e/a0691e00.htm>, from Joint FAO/WHO Expert Committee on Food Additives (JECFA) <http://www.fao.org/food/food-safety-quality/scientific-advice/jecfa/jecfa-additives/en/>

- Food and Agriculture Organization. Specifications for Flavours. 2006b. (online database & report). Retrieved <http://www.fao.org/food/food-safety-quality/scientific-advice/jecfa/jecfa-additives/en/> from Joint FAO/WHO Expert Committee on Food Additives (JECFA) <http://www.fao.org/food/food-safety-quality/scientific-advice/jecfa/jecfa-flav/en/>
- Food and Drug Administration. FDA analysis finds bread grossly mislabeled: Enforcement Story 2003: Center for Food Safety and Applied Nutrition. 2003. Retrieved August 8, 2014, from <http://www.fda.gov/iceci/enforcementactions/enforcementstory/enforcementstoryarchive/ucm095929.htm>
- Guidance for Industry: A Food Labeling Guide (6. Ingredient Lists). 2013
- Food and Drug Administration. Proposed Rule. Department of Health and Human Services, FDA; 2014a. Food Labeling: Revision of the Nutrition and Supplement Facts Labels. A Proposed Rule by the Food and Drug Administration on 03/03/2014; p. 11879-11987. March 3, 2014 ed., Vol. Docket No. FDA-2012-N-121011109 pages
- Food and Drug Administration. Labeling & Nutrition: Food Labeling and Nutrition Overview. 2014b. Retrieved February 4, 2011, from <http://www.fda.gov/Food/LabelingNutrition/default.htm>
- Johnson RK, Appel LJ, Brands M, Howard BV, Lefevre M, Lustig RH, Wylie-Rosett J. Dietary sugars intake and cardiovascular health: a scientific statement from the American Heart Association. [Practice Guideline]. *Circulation*. 2009; 120(11):1011–1020. doi:10.1161/CIRCULATIONAHA.109.192627. [PubMed: 19704096]
- Johnson RK, Appel LJ, Brands M, Howard BV, Lefevre M, Lustig RH. Dietary Sugars Intake and Cardiovascular Health: A Scientific Statement From the American Heart Association. *Circulation*. 2009; 120(11):1011–1020. Prevention. doi: 10.1161/circulationaha.109.192627. [PubMed: 19704096]
- Malik VS, Popkin BM, Bray GA, Despres JP, Willett WC, Hu FB. Sugar-sweetened beverages and risk of metabolic syndrome and type 2 diabetes: a meta-analysis. [Meta-Analysis Review]. *Diabetes Care*. 2010; 33(11):2477–2483. doi: 10.2337/dc10-1079. [PubMed: 20693348]
- Microsoft. (2013). Microsoft Excel (Windows). Seattle WA, Morenga LAT, Howatson AJ, Jones RM, Mann J. Dietary sugars and cardiometabolic risk: systematic review and meta-analyses of randomized controlled trials of the effects on blood pressure and lipids. *Am J Clin Nutr*. 2014; 100(1):65–79. doi: 10.3945/ajcn.113.081521. [PubMed: 24808490]
- Murphy, S. Integrity checks for nutrient data; Paper presented at the Proceedings of the Fourteenth National Nutrient Databank Conference; Iowa city, IA. 1989;
- National Archives and Records Administration. 21 CFR 101.9. Washington, DC: 1993. Retrieved from <http://ecfr.gpoaccess.gov/cgi/t/text/text-idx?c=ecfr&sid=563f0b6235da3f4c7912a64cbceec305&rgn=div8&view=text&node=21:2.0.1.1.2.1.1.6&idno=21>
- NIST Handbook 133 Checking the Net Contents of Packaged Goods. 2005
- Ng SW, Popkin BM. Monitoring Foods and Nutrients Sold and Consumed in the United States: Dynamics and Challenges. *Journal of the Academy of Nutrition and Dietetics*. 2012; 112(1):41–45.e44. doi:<http://dx.doi.org/10.1016/j.jada.2011.09.015>. [PubMed: 22389873]
- Ng SW, Slining MM, Popkin BM. Use of Caloric and Noncaloric Sweeteners in US Consumer Packaged Foods, 2005-2009. *J Acad Nutr Diet*. 2012; 112(11):1828–1834.e1826. [PubMed: 23102182]
- Ockerman, HW. Food Science Sourcebook. Vol. 1 & 2. Van Nostrand Reinhold; New York City: 1991.
- Reedy J, Krebs-Smith Susan. Dietary Sources of Energy, Solid Fats, and Added Sugars among Children and Adolescents in the United States. *Journal of the American Dietetic Association*. 2010; 110(10):1477–1484. [PubMed: 20869486]
- SAS Institute Inc.. SAS (c) 9.3. Cary, NC: 2011.
- SAS Institute Inc.. SAS/IML (c) 12.1 [SAS Analytic Products]. Cary, NC: 2012.
- Schakel SF, Buzzard IM, Gebhardt SE. Procedures for Estimating Nutrient Values for Food Composition Databases. *Journal of Food Composition and Analysis*. 1997; 10(2):102–114. doi: <http://dx.doi.org/10.1006/jfca.1997.0527>.

- Slining M, Yoon EF, Davis J, Hollingsworth B, Miles D, Ng SW. An approach to monitor food and nutrition from 'Factory to Fork.' J Acad Nutr Diet. 2015; 115(1):40–49. doi: doi:10.1016/j.jand.2014.09.002. [PubMed: 25441958]
- Te Morenga L, Mallard S, Mann J. Dietary sugars and body weight: systematic review and meta-analyses of randomised controlled trials and cohort studies. [Journal Article]. British Medical Journal. 2013; 346:e7492. doi: 10.1136/bmj.e7492. [PubMed: 23321486]
- U.S. Department of Agriculture. Report of the Dietary Guidelines Advisory Committee on the Dietary Guidelines for Americans, 2010. 2010. Available at: <http://www.cnpp.usda.gov/DGAs2010-DGACReport.htm>
- U.S. Department of Agriculture. U.S. Department of Health and Human Services. Dietary Guidelines for Americans, 2010. U.S. Government Printing Office; Washington, DC: 2010.
- Urban LE, Dallal GE, Robinson LM, Ausman LM, Saltzman E, Roberts SB. The Accuracy of Stated Energy Contents of Reduced-Energy, Commercially Prepared Foods. Journal of the American Dietetic Association. 2010; 110(1):116–123. doi: <http://dx.doi.org/10.1016/j.jada.2009.10.003>. [PubMed: 20102837]
- Welsh JA, Sharma A, Cunningham SA, Vos MB. Consumption of Added Sugars and Indicators of Cardiovascular Disease Risk Among US Adolescents. Circulation. 2011; 123(3):249–257. doi: 10.1161/circulationaha.110.972166. [PubMed: 21220734]
- Westrich BJ, Altmann MA, Potthoff SJ. Minnesota's Nutrition Coordinating Center Uses Mathematical Optimization to Estimate Food Nutrient Values. Interfaces. 1998; 28(5):86–99.
- Westrich BJ, Buzzard IM, Gatewood LC, McGovern PG. Accuracy and Efficiency of Estimating Nutrient Values in Commercial Food Products Using Mathematical Optimization. Journal of Food Composition and Analysis. 1994; 7(4):223–239.

Highlights

- We developed a batch approach to estimate added sugars in Consumer Packaged Goods
- We estimated added sugars content of 7021 beverages with any sugar
- Results show it is possible to derive valid estimates of added sugars in CPGs
- We can monitor the presence and amount of added sugars in the food supply/purchases

Figure 1. Overview of batch-mode approach to estimate added sugars in CPGs

Notes: CPG= Consumer Packaged Goods; NFL= Nutrition Facts Label

Figure 2. Simple peanut butter example of how Linear Programming (LP) model works

Notes: LP= Linear Programming; NFL= Nutrition Facts Label

Figure 3. Contribution of added vs. intrinsic sugars to total sugars in ready-to-drink CPG beverages, 2007-08

Notes: CPG= Consumer Packaged Goods; UPC= Universal Product Code; NFL= Nutrition Facts Label Figure only includes U PCs with NFL total sugars >0g. Some categories contain both regular and low calorie products. Fruit juice and fruit juice concentrate not considered an added sugar.

Table 1
Nutrient and Weight ‘errors’ that are equivalent to 1 tolerance unit (as applied to this study)

	Per 100 g
Weight tolerance	±5 g
Nutrient tolerance *	
Calories	±85 kcal
Total fat	±2.5 g
Saturated fat	±2.5 g
Total cholesterol	±30 mg
Sodium	±100 mg
Total carbohydrate	±10 g
Total sugars	±10 g
Fiber	±2.5 g
Protein	±5 g
Potassium	±100 mg

* Source: Nutrition Coordinating Center, University of Minnesota [Table 1 of Westrich (1994)].

Table 2

Macronutrient values (LP estimated to NFL) and added sugar estimates across beverage categories

Beverage category	Caloric sodas & energy drinks			Sports drinks			Flavored water – carbonated & still			Fruit & vegetable juice drinks			Fruit flavored drinks			Sweetened dairy & dairy alternatives			Ready-to-drink coffees & teas			Concentrated fruit drinks			Powdered beverage mixes			Cocktail mixes		
	N	mean	SD	N	mean	SD	n	mean	SD	n	mean	SD	N	mean	SD	N	Mean	SD	N	mean	SD	n	mean	SD	n	mean	SD	n	mean	SD
# of unique products in 2007-08																														
with NFL total sugar >0g	1711			290			1649			183			860			742						130			942			183		
with estimated added sugar >0g	1695			290			1524			183			829			712						124			901			174		
LP to NFL Nutrient Ratios																														
Calories	1711	1.03	0.07	290	1.08	0.14	1652	1.04	0.20	1652	1.01	0.16	183	1.06	0.69	867	1.01	0.13	769	1.02	0.18	130	1.00	0.08	942	0.99	0.29	184	1.00	0.09
Total fat	7	0.86	0.37	0			44	0.78	0.40	44	1.98	2.21	1	0.02		836	1.05	0.31	140	0.96	0.32	7	0.97	0.46	493	1.10	0.33	18	0.59	0.47
Saturated fat	2	1.36	0.64	0			22	1.16	0.23	22	3.31	2.77	0			749	0.96	0.25	131	0.92	0.30	5	1.01	0.07	389	0.89	0.39	10	0.71	0.42
Carbohydrate	1711	1.00	0.04	290	1.01	0.07	1652	1.01	0.06	1652	1.01	0.10	183	1.08	0.75	867	1.01	0.15	769	1.00	0.16	130	1.01	0.07	942	1.01	0.28	184	1.01	0.08
Total sugar	1711	0.99	0.05	290	0.99	0.07	1649	0.97	0.22	1649	1.00	0.24	183	1.04	0.70	860	0.93	0.19	742	0.99	0.19	130	1.01	0.08	942	1.00	0.16	183	1.02	0.19
Protein	61	0.19	0.35	1	1.00		220	0.98	0.47	220	0.73	0.65	4	0.00	0.00	849	1.06	0.27	152	0.89	0.39	50	0.23	0.32	504	1.03	0.28	28	1.08	0.74
Est. added sugar (g per 100g of product) among UPCs with NFL total sugar >0g	1711	10.79	2.31	290	5.21	1.63	1649	4.31	3.23	1649	8.00	3.84	183	8.03	4.90	860	6.27	3.61	742	7.44	4.77	130	36.11	14.41	942	66.17	26.23	183	19.36	11.20
Est. added sugar calories as % of calories among UPCs with NFL total sugar >0g	1711	98.25%		290	101.46%		1649	91.54%		1649	71.06%		183	97.53%		860	33.80%		742	79.62%		130	77.64%		942	67.73%		183	84.40%	
Median total error per nutrient and weight (tolerance units)		0.008			0.006			0.003			0.010			0.004			0.034			0.008			0.065			0.136			0.012	

Notes: LP= Linear programming; NFL= Nutrition Facts Label; UPC= Universal Product Code; SD= Standard deviation. Estimated values from LP model when fruit juice and fruit juice concentrate not considered an added sugar. Some beverage categories contain both regular and low calorie products. Ns represent products with non-missing or non-zero NFP values.

Table 3
Validation test results for four beverages with different formulations

A. Chocolate Milk Ingredient (per 100g of product)	Formula 1		Formula 2	
	Actual (g)	LP (g)	Actual (g)	LP (g)
1% Lowfat Milk	95.65	93.81	93.15	91.67
Sugar [‡]	2.50	2.54	5.00	5.12
Alkalized cocoa	1.50	1.10	1.50	0.85
Salt	0.15	0.15	0.15	0.13
Carrageenan	0.10	0.07	0.10	0.23
Artificial Chocolate Flavor	0.10	2.00	0.10	2.00
Vitamin A Palmitate	0.0008	0.00	0.0008	0.00
Vitamin D3	0.00004	0.33	0.00004	0.00
Added Sugar	2.50	2.54	5.00	5.12
Total Sugar	7.30	7.30	9.64	9.64

B. Cranberry Grape Drink Ingredient (per 100g of product)	Formula 1		Formula 2	
	Actual (g)	LP (g)	Actual (g)	LP (g)
Water	74.19	75.91	50.69	50.34
Grape Juice from concentrate	12.50	11.55	40.50	32.88
Sugar [‡]	10.50	10.40	6.00	6.32
Cranberry Juice from concentrate	2.50	2.00	2.50	6.32
Black Carrot Juice (for color)	0.155	0.00	0.155	2.00
Sodium Citrate	0.13	0.14	0.13	0.14
Ascorbic Acid	0.025	0.00	0.025	2.00
Added Sugar	10.50	10.39	6.00	6.32
Total Sugar	12.46	12.46	12.09	12.31

C. Eggnog Ingredient (per 100g of product)	Formula 1		Formula 2	
	Actual (g)	LP (g)	Actual (g)	LP (g)
Milk	63.00	63.53	50.00	49.01
Sugar [‡]	9.00	10.82	18.00	18.72
Cream	9.00	10.82	12.00	14.51
Skim Milk	8.80	4.70	7.80	4.84
Corn Syrup [‡]	5.00	4.70	7.00	4.84
Egg yolks	5.00	4.70	5.00	4.84
Water	0.1092	0.02	0.1092	2.00
Carrageenan	0.05	0.16	0.05	0.15
Nutmeg	0.04	0.55	0.04	0.00
Vitamin A Palmitate	0.0008	0.00	0.0008	1.09
Added Sugar	11.53	13.19	21.54	21.15
Total Sugar	16.58	16.58	26.51	24.04

D. Vegetable juice drink Ingredient (per 100g of product)	Formula 1		Formula 2	
	Actual (g)	LP (g)	Actual (g)	LP (g)
Water	44.405	34.43	76.385	71.83
Tomato Juice from concentrate	40.00	34.43	10.00	11.27
Carrot Juice from concentrate	10.00	26.27	7.00	10.30
Sugar [‡]	5.50	4.87	6.52	6.21
Citric Acid	0.04	0.00	0.04	0.00
Malic Acid	0.03	0.00	0.03	0.39
Ascorbic Acid	0.025	0.00	0.025	0.00
Added Sugar	5.50	4.87	6.52	6.21
Total Sugar	7.37	7.02	6.98	6.98

Notes: LP= Linear programming

[‡] denotes defined as an added sugar; fruit juice and fruit juice concentrate not considered an added sugar. Added sugar (g) values only include the proportion of added sugar ingredients that are sugar. For all formulations we assumed that the ingredient lists included 'manufacturer <2%' statements

Table 4
Estimates on number of products that contain added sugar and their added sugar content
among products that contain any sugar in select beverage categories

Added Sugar definition	FJC and FJ not considered added sugar	FJC (but not FJ) considered added sugar	Both FJC and FJ considered added sugar
<i>Caloric sodas & energy drinks</i>			
N with total sugar >0g	1711	1711	1711
N with added sugar >0g	1695	1699	1709
Added sugar g/100g among products with total sugar >0g, mean (SD)	10.79 (2.31)	1.04 (2.04)	11.09 (1.97)
Added sugar calories as % of total sugar calories, %	98.25%	100%	100%
<i>Flavored waters – carbonated & still</i>			
N with total sugar >0g	331	331	331
N with added sugar >0g	297	319	327
Added sugar g/100g among products with total sugar >0g, mean (SD)	4.31 (3.23)	4.47 (3.13)	4.47 (3.13)
Added sugar calories as % of total sugar calories, %	91.54%	94.83%	94.90%
<i>Fruit and vegetable juice drinks</i>			
N with total sugar >0g	1649	1649	1649
N with added sugar >0g	1524	1595	1644
Added sugar g/100g among products with total sugar >0g, mean (SD)	8.00 (3.84)	9.25 (3.63)	10.20 (3.14)
Added sugar calories as % of total sugar calories, %	71.08%	82.18%	90.65%
<i>Fruit flavored drinks</i>			
N with total sugar >0g	183	183	183
N with added sugar >0g	183	183	183
Added sugar g/100g among products with total sugar >0g, mean (SD)	8.03 (4.90)	8.09 (4.86)	8.09 (4.86)
Added sugar calories as % of total sugar calories, %	97.53%	98.24%	98.24%
<i>Sweetened dairy and dairy alternatives</i>			
N with total sugar >0g	860	860	860
N with added sugar >0g	829	832	832
Added sugar g/100g among products with total sugar >0g, mean (SD)	6.27 (3.61)	6.28 (3.61)	6.28 (3.61)
Added sugar calories as % of total sugar calories, %	33.80%	33.87%	33.87%

Notes: FJC= Fruit Juice Concentrate; FJ= Fruit Juice; SD= Standard deviation