Coding Recipes: Dilemmas and Decisions

Betty Perloff, Human Nutrition Information Service,
U.S. Department of Agriculture

Calculating the nutrient content of recipes based on data for ingredients is considered standard operating procedure for the nutritional analysis of dietary intake data. Nutrient calculations for recipes undoubtedly will continue until quick and inexpensive, as well as reliable, laboratory methods of analysis have been developed. Since those analytical procedures do not appear to be on the horizon, we must strive to perfect the recipe calculation procedures to the extent possible.

Kristin Marcoe has described the recipe calculation procedure we use at USDA's Human Nutrition Information Service (HNIS). Many decisions were made during the process of automating this procedure, and many more are made in the day-to-day operation of the system. Today I will present examples of some of the choices we have faced and how the decisions have been approached.

When we began planning the automation of recipe calculations for the Survey Nutrient Data Base, the first decision we faced was to choose an appropriate method for the calculations. Our goals were to select a method that (1) was research based, (2) was shown to be most comparable to analytical results, and (3) that could be updated as new research provides us with better information about changes that take place during cooking or food processing.

In addition to the retention factor method we eventually chose, we also considered two other procedures. The first alternative was also a retention factor method, but it differed from our final selection in that it applied one set of retention factors to the recipe instead of different factors to each ingredient. We referred to this alternative method as the "dish" retention method, as opposed to the "ingredient" retention method that was finally chosen.

The second alternative was to use yield factors for the individual ingredients, converting them to the weight we expected after cooking and applying the nutrient values for the cooked item.

We sponsored two research projects with Oregon State University to provide data for individual ingredients and for recipes cooked from those ingredients. Information from those projects, along with other available data on yields and retentions, were then used to study applications of the three recipe calculation methods.

From this project, we were able to ascertain that the two retention factor methods gave comparable results. Otherwise, the results were largely inconclusive; however, we were able to develop a set of pros and cons for each method. The resulting recommendation was to use the "dish" retention method, and the computer program was originally written for that method. The original decision was based on several factors. Primarily, the reasons were (1) The "dish" retention method was considered the traditional method, since it had been used for earlier editions of Agriculture Handbook No. 8; (2) it would allow for interactions among foods that might affect retention of nutrients; and (3) the method appeared simpler than the ingredient retention method.

However, when we applied this method to our daily work we found it was not simpler to use, because existing retention data were primarily for individual ingredients, not complete mixtures. Using this method frequently required us to calculate the dish retention factors based on the ingredient retention factors. In effect, the process had been complicated, not simplified. Furthermore, prospects for obtaining adequate numbers of "dish" retention factors through research contracts was dim because of the many different types of mixtures that were appearing in our food consumption surveys. If the need to allow for interactions of foods during cooking were to arise, we realized it could be compensated for by additional ingredient factors taking into account other types of foods that might be present.

We quickly revised our computer program to accommodate the ingredient retention factors and computerized the table of percent retention of nutrients in food preparation. As new retention data have become available, we have revised and expanded the table of retention factors.

Recipe Selection

Other types of recipe-related decisions are made on a regular basis at HNIS. The most frequent decision that is faced is the selection of the recipe. Recipe information is frequently provided by survey respondents for home-prepared foods, although this information is seldom complete. For example, recipes for homemade soups usually have no mention of the amount of water or other liquid used in preparation. For foods eaten away from home, the main ingredients or characteristics of the recipe are usually the extent of reported information.

When different versions of recipes already existing in our system are encountered, we now have the ability to modify the existing recipes. So far, this recipe modification feature has been used only in the pilot test for our next survey. Its primary use has been to allow for different types of fats used in recipes or to change or add other ingredients that would likely have a dramatic effect on a recipe's nutrient content. The results of this pilot test appear promising, and we are cautiously optimistic about the potential for this new feature to provide us with greater flexibility for capturing more specific recipe information from our survey respondents.

When new or unique recipes are encountered, the first step is to locate the same or similar item in recipe books. We maintain a supply of current popular cookbooks, as well as selected regional and specialty cookbooks. We try to locate recipes from a minimum of three sources, looking first at the popular cookbooks. Those recipes are compared with each other and with the respondent's recipe. If the same recipe is found in at least two of the sources, it is selected. However, if only widely varied recipes are located, a composite recipe may be constructed. Recipes are reviewed periodically and revised when warranted. Frequently consumed items receive priority for review.

Recipes for commercially prepared mixtures may represent more than one brand name. When a new brand is reported, ingredient labels are compared to the existing recipe. If they are similar, the brand name is added to the food item's description. If they are different, a new formulation estimate is prepared based on the list of ingredients and any nutrient data that are available.

Ingredient selections

Once a recipe is selected, many decisions still remain. Ingredients are matched to identical items on the data base where they exist. For home-prepared foods, ingredients are usually matched to the form of the item that is identified in the recipe. Fresh items are assumed if other forms are not designated. When cooking is applied to an ingredient prior to incorporating it into the recipe, a yield factor is applied to the weight and the cooked form of the food is selected from the data base. For example, if a recipe calls for 1 pound of macaroni prepared according to package directions, then the weight associated with the yield from 1 pound of macaroni after cooking, 1,140 grams, is used with the data base item for the cooked form of macaroni. Likewise, "1 pound of ground beef, browned," is translated into the cooked weight, 352 grams, and used with the appropriate cooked data base item.

When ingredients are missing from the data base, a closely related item is substituted. Missing ingredients for which substitutions are required are flagged, enabling us to track the frequency of their use and to include them in plans for analytical research when appropriate. For example, when we needed to expand the data base for Mexican-American foods reported in Hispanic HANES, several recipes called for Mexican cheeses that did not exist in the data base. We matched them as closely as possible to existing items, and then targeted the Mexican cheeses for analysis when new research was planned.

Commercially prepared foods may include ingredients not found in the data base. We've added some special items, cellulose for example, to facilitate formulating commercial items. We've also added special data records to represent added ascorbic acid and added calcium in commercially prepared foods. Ingredients for commercially prepared frozen entrees are usually matched to frozen forms of the items.

Ingredient weights

Weights for ingredients are selected based on the description of the measure. We maintain an extensive data base of measure descriptions and weights. We have a Weights and Measures Team that is responsible for identifying discrepancies that arise in weights for various measure descriptions, and we are fortunate to have a modern, well-equipped laboratory where discrepancies can be resolved and where weights for new foods can be determined.

Retention codes

To apply retention factors to ingredients for estimating cooking losses, the ingredients are matched against our retention factors description file. This file contains descriptions of various categories of foods and cooking methods for which retention factors are available. Sometimes cooking times are also included in the retention factor descriptions, but specific cooking times are frequently not available. For example, a recipe may call for heating milk 20 minutes; however, the choices for available retention factors are 10 minutes, 30 minutes, and 1 hour. In this example, we chose the retention factors for 30 minutes, providing the more conservative estimate of nutrient content.

Retentions are not always present for the specific cooking method either. Again, we must decide the most similar method. For example, when coding the recipe for doughnuts, the available retention factor categories for flour were baked, boiled, reheated, sautéed, and toasted. Sautéed was selected in this case.

Yield factors

Selecting appropriate yield factors to represent the changes that take place in moisture and fat content of foods during cooking frequently pose difficult choices, and this is an area that may be in greatest need for additional research. The major source of this type of information is Agriculture Handbook No. 102, "Food Yields After Different Stages of Preparation." We match newly coded recipes against previously coded ones, selecting the closest match for type of recipe, ingredients, and cooking method. For example, when we coded the recipe for moussaka, an eggplant and meat casserole, we matched it to tuna noodle casserole and estimated the loss in weight to be 10 percent.

Comparisons—calculations versus analyses

We're frequently asked how accurate are the nutrient values generated from recipe calculations. Obviously, they can be no better than the research on which they are based. Part of the purpose of the original research we sponsored before designing our recipe calculation program was to answer that question. Results from that research were reported at past Nutrient Data Bank Conference. Calculated values for proximate components and minerals usually fell within 10 percent of the analyzed values. The differences for copper, however, were higher. Calculated values for vitamins were usually within 20 percent of the analyzed values; however, differences for vitamin B-12 were greater.

We have made many improvements to our data base of nutrient values for ingredient items since those comparisons were reported. Additional refinements to yield and retention data are needed and may also improve recipe calculation results.

Conclusions

Calculations to determine the nutrient content of recipes are used extensively for the nutritional analysis of dietary intake data at the USDA's Human Nutrition Information Service. Nutrient estimates derived through these calculations are quick and inexpensive, and they serve a very useful purpose. However, many decisions are required before adequate values can be calculated. HNIS selected a recipe calculation method based on research and practical considerations. Daily operations require decisions on recipe selection as well as on ingredients, weights, and retention and yield factors.