Issues of Food Description

by Jean A.T. Pennington, Food and Drug Administration, Washington, DC

Food names and descriptions identify and distinguish among foods listed in databases. Descriptive terms may be captured in free form with the food name and in faceted systems. The INFOODS Food Description System is faceted with free text. The International Interface (which includes LANGUAL) is a faceted system with standardized vocabulary. Faceted systems provide a checklist to fully describe foods and they allow for retrievals and matching of foods among databases. On-line dictionaries can be used to clarify implicit and complex food names. Relevant descriptive information varies among foods and food types. Information about cooking methods, ingredients, recipes, cuisine, and preparation location are important to fully define some foods. Sample descriptions (as opposed to food descriptions) identify the products analyzed in the laboratory. Market share and default entries should be clearly identified. Pictures of foods (hardcopy or computerized) are useful for product identification. A universal system to describe foods will enhance the sharing and exchange of food composition data.

Introduction

Clear, accurate food descriptions are essential to identify and distinguish among foods listed in databases. Such descriptions enable database users to select the most relevant foods, obtain the most appropriate nutrient data, and avoid mistakes that might occur by selecting the wrong products. The number and types of food products available to consumers is huge and continues to grow. Productive, competitive food companies continue to challenge the retail market with new products. Trade across national borders enhances the availability of foreign food products. Even those who are experienced in the development and use of food composition databases encounter food names that are new or unusual. It is a continuing challenge to keep national databases current with the available food supply and to describe foods in a manner that is consistent and useful. This paper identifies and discusses issues relative to food names and descriptive terms used in databases and offers some suggestions to improve and enhance food descriptions.

How to Capture Descriptive Information

Descriptive information for foods in databases has traditionally been captured in free form (free text) with the food name. The information is usually recorded when the food is obtained for laboratory analysis. Relevant descriptive information distinguishes one product from another and may affect nutrient values. Most databases impose some order on descriptive terms to facilitate use of the database.

Descriptive information may also be captured in a specific format using a faceted system. The INFOODS Food Description System (Truswell et al., 1991) is a hardcopy, faceted system with free text. LANGUAL, a computerized food description vocabulary developed at the United States (US) Food and Drug Administration (FDA) beginning in 1975, uses standardized terms for 15 descriptive facets of foods (McCann et al., 1988). One of the unique features of LANGUAL is the hierarchy of standardized terms for each facet. The most relevant terms are selected to describe

each food. Under contract with Technical Assessment Systems in Washington, DC, FDA developed an International Interface Standard which includes LANGUAL plus other important aspects of food description (Pennington & Hendricks, 1992, 1995).

Implicit/Complex Food Names

Implicit food names are unfamiliar names that have no meaning without prior knowledge or experience with the foods. An example is "hush puppies" (deep-fried cornmeal bread), a food of southern cuisine in the United States (US). Some food names are developed by industry and have no inherent meaning such as Whopper (fast-food hamburger) and Frankenberry (breakfast cereal). Some food names are lengthy as well as unclear or implicit such as "Ben & Jerry's Chocolate Chip Cookie Dough Ice Cream" and "I Can't Believe It's Not Butter."
Dictionaries for databases can define implicit and complex food names and identify preferred food names and synonyms. Database dictionaries are useful when national databases are used in other countries. The dictionaries can be on-line with computerized databases so that definitions of food names can be seen along with the nutrient data for foods.

Relevant Information

The way foods are described in databases is usually related to the intended uses of the databases. Databases developed for dietary assessment describe foods in the way that survey participants describe them. This makes it easier to find the best match between a food eaten and one listed in a database. Food databases used for other purposes (e.g., institutional menu development and inventory, industry product development, or epidemiology studies) have different emphases for food descriptions related to cost, technology, health issues, nutrient levels, or ingredients. Relevant descriptive information also varies among foods and food types. For a raw, agricultural product (e.g., raw apple), the relevant information may relate to color, cultivar, Latin name, geographical location, season, year of harvest, soil type, use of pesticides or herbicides, storage time, and part consumed. For a processed or restaurant food such as frozen macaroni and cheese or fast-food hamburgers, the relevant information may be the brand name, the name and location of the manufacturer or restaurant, the ingredients and recipe, storage time, weight of the serving, production date, and preparation instructions. For a homemade food such as lasagna, the relevant information might be geographical area, cuisine, ingredients, and recipe. The Interface Standard allows for capture of these different facets using specific descriptive terms. Facets may be left blank if they are not relevant to a food or the information is not known. The terms "not known" and "not relevant" may be entered for clarification.

Cooking Methods

Specific information is needed about the methods used to cook foods because the methods can affect nutrient values and may be of interest regarding health matters (e.g., grilling, microwave use). Cooking methods that add fat (frying, sautéing) or cause loss of water (boiling, frying) or fat (broiling) lead to changes in nutrient values. Cooking method is included in the Interface Standards as one of the LANGUAL facets.

Ingredients and Recipes

Recipes specify the quantities of ingredients and explain how they are put together to form the completed product. This information is especially important for mixed dish and multi-ingredient foods that have the same names but contain different ingredients or are put together in different ways. Information about ingredients and recipes helps clarify the differences in nutrient values and assists data users in selecting the correct items from the database. Ingredients and recipes have been included as part of the International Interface Standard and are available on-line as one works with a database.

Cuisine

Populations in all countries are increasingly exposed to foods of different cultures and cuisines. Ethnic and cultural foods are widely distributed and available in markets, restaurants, and ethnic food stores. The mobility of populations also enhances food experiences. Information about cuisine in databases is important to help identify foods. It is also important to help distinguish foods from different cuisines that have the same names. For example, "tuna" is a fish in the US and a prickly pear in Mexico. "Tortilla" is a flat, unleavened pancake of corn or wheat flour in Mexico and the US, but is an omelet in Spain.

A draft cuisine hierarchy was developed for the Interface Standard to identify the cultural background of foods. Terms for the hierarchy were based on restaurant designations (in travel guides and phone books), culinary schools, scientific and cultural literature, and cookbooks. Eight groups of cuisine origin were identified, African, Caribbean, East Asian, European, Latin American, Mid-Eastern, Native American, and Native Australian. Among these eight general groups, 150 unique cuisines were identified.

Cuisine terms may be used in a national database to identify foods that are not part of the national cuisine. When ethnic foods (or foods that are different from a national cuisine) are entered into a database, the database compiler must decide whether to use the original food names or translations of the food names as the preferred terms. Databases used for national surveys in the US, have original food names for many Mexican, Asian, and Italian foods as preferred entry terms, but have American translations as the preferred entry terms for foods that are not well known, e.g., rice with chicken, Puerto Rican style (arroz con pollo).

Where Foods are Prepared/Obtained

Homemade, restaurant, and commercially manufactured versions of a food may have different nutrient content. Unfortunately, many databases do not specify the source of the food or location of preparation. The Interface Standard allows for information on the location of food preparation. Information about ingredients and recipes also helps distinguish between homemade, restaurant, and commercially manufactured foods.

Excess Information

The foods described in databases should match the product analyzed for nutrients in the laboratory. However, some descriptive information may not affect nutrient values, and database compilers usually choose to have single entries for foods that differ only in minor ways such as shape (e.g., pineapple slices or chunks) or flavor (e.g., plain or onion flavored potato chips). It is best to merge

entries in databases based on nutrient content rather than have multiple listings of essentially similar foods. Extensive listings of similar foods may confuse and frustrate database users.

Sample Description, Market Share, Default Entries

The distinction between descriptive information for a food and descriptive information for the analytical sample is not always clear. Both types of descriptions are important to fully understand what the food data represent. Sample descriptions usually specify the variables included in the sampling plan. For example, the food description "Plaice, fried in retail blend oil" is accompanied by the sample description "20 samples purchased from fish and chip shops" in the British food composition table (Holland et al., 1993). Other examples are "Herring, canned in tomato sauce; 10 samples, 4 brands, whole contents" and "Apples, cooking, baked w/o sugar, flesh and skin; 10 samples, cored and baked, 180 C, 30-40 minutes" (Holland et al., 1992, 1993).

The nutrient values for market share entries in databases are determined by weighting nutrient data from different cultivars, seasons, brands, etc. The data are weighted based on the availability of a food to a defined population during a specific time period. Market share entries are useful when general nutrient information is needed such as average, year-round values for grapefruit or average values for various brands of enriched white bread. Market share data should be clearly identified in databases, e.g., grapefruit, raw (US market share, 1995); enriched white bread (US market share, first quarter 1993). The market share for grapefruit in the US would reflect a specific percent of pink and red and white grapefruit from California, Florida, and Texas, while the market share for enriched white bread would reflect sales of major brands.

Default entries are useful for survey databases when subjects in dietary studies or surveys cannot accurately describe the foods they ate. The databases for national food consumption surveys in the US have "not further specified" (NFS) entries such as meat, NFS; sandwich, NFS; and fruit pie, NFS. The data for these products may be based on analytical data for the most frequently consumed products (e.g., data for frozen apple pie may be used for pie, NFS) or calculated based on market shares (e.g., hamburger, NFS might be based on market shares of hamburgers from the top four fast-food chains).

Pictures - Hardcopy and Computerized

Color pictures in hardcopy and computerized databases allow foods to be more easily and specifically identified. They are especially useful for unfamiliar foods. Pictures are helpful to participants of food consumption surveys as they try to remember and name the foods they consumed.

Pictures are particularly useful for basic foods (fruits, vegetables, nuts, eggs, legumes) and for some pastries and breads, mixed dishes (tostadas, open-faced sandwiches, salads), and commercial products. Pictures of commercial products may display the packaged items. Pictures show differences between countries and regions in meat cuts and the degree of fat trim, and they distinguish between the parts of plants that are consumed and the parts discarded.

Pictures may not be useful for layered recipe items if the top layers are different from the lower layers (e.g., tuna noodle casserole topped with cheese and bread crumbs, beef and cheese enchiladas). They may also be confusing if there are various presentations of the same product, i.e., foods with the same name, but different appearances in pictures.

Advantages of Faceted Systems

Faceted systems provide a checklist as a reminder to fully describe foods. They also allow for retrievals of foods from databases based on standardized, descriptive terms. All food names in a database that have a particular characteristic or mixture of characteristics can be identified and retrieved. For example, one may want to retrieve all foods of Chinese cuisine that contain beef or all frozen foods in paperboard containers. The computer searches for the standardized terms and provides a printout of the identified foods, saving the users from tedious, and perhaps incomplete, manual searches. The retrieval feature is available on the Interface Standard.

Faceted systems may also be used to find the best match for a food in a database. The initial food may be one described by a survey participant, one analyzed in a laboratory, or one from another database. To match a food, the computer retrieves from the database all foods with the same (or almost the same) descriptive terms.

The use of faceted systems should not decrease the descriptive information presented with the food name. Some descriptive information should stay with a food name in a database, some should go into computerized facets (or hardcopy background documentation), and some should go in both places. Information regarding color, maturity, cultivar, part used, preparation state, physical state, preservation method, cooking method, and brand name should be included with the food name in databases and also in facets if a faceted system is used. If the food was obtained from a fast-food or other restaurant or is homemade, and this is not clear from the name, the term "fast-food," "restaurant," or "homemade" should be added. If the cuisine is regional or otherwise different from the national cuisine, the cuisine term should be included with the food name. Information about the language of the food name, the geographical location of where the food was obtained, and agricultural conditions can be included in the facets. If the food groupings are not otherwise indicated, they can be added if there is a possibility of confusion (e.g., chili as vegetable, powder/spice, or mixed dish; tuna as fish or vegetable).

Conclusions

There are currently a number of barriers to the sharing and exchanging of data among food composition databases. These barriers include differences in nutrient definitions, analytical methods, data quality, units, and food descriptions. The move toward a universal system to describe foods is a desirable one that will allow database users to more readily share and exchange food composition data. Different and innovative ways of describing and accessing food data should be considered to determine what works best. A system that merges the best aspects of current systems should be devised. Until a universal system is agreed upon, database developers should continue to promote the capture of thorough and accurate descriptive information about foods.

References

Holland, B., Unwin, I.D. & Buss, D.H. (1992). *Fruit and Nuts*, First Supplement to the 5th ed. of McCance and Widdowson's The Composition of Foods, The Royal Society of Chemistry, Cambridge.

Holland, B., Brown, J. & Buss, D.H. (1993). *Fish and Fish Products*, Third Supplement to the 5th ed. of McCance &Widdowson's The Composition of Foods, The Royal Society of Chemistry, Cambridge.

McCann, A., Pennington, J.A.T., Smith E.C., Holden, J.M., Soergel, D. & Wiley, R.C. (1988). FDA's factored food vocabulary for food product description. *J. Am. Diet. Assoc.* 88, 336-341.

Pennington, J.A.T. & Hendricks, T.C. (1992). Proposal for an international interface standard for food databases. *Food Addit. Contam.* 9, 265-275.

Pennington, J.A.T., Hendricks, T.C., Douglass, J.S., Petersen, B., Kidwell, J. (1995) International interface standard for food databases. *Food Addit. Contam.* in press.

Truswell, A.S., Bateson,, D.J., Madafiglio, K.C., Pennington, J.A.T., Rand, W.M. & Klensin, J.C. (1991). Committee report: INFOODS guidelines for describing foods: a systematic approach to describing foods to facilitate international exchange of food composition data. *J. Food Comp. Analysis* 4, 18-38.