



Original Article

Extending a multiethnic food composition table to include standardized food group servings

S. Sharma*, S.P. Murphy, L.R. Wilkens, D. Au, L. Shen, L.N. Kolonel

Etiology Program, Cancer Research Center of Hawaii, University of Hawaii, 1236 Lauhala Street, Honolulu, HA 96813, USA

Received 22 May 2002; received in revised form 5 December 2002; accepted 9 December 2002

Abstract

There is a growing awareness that foods may have health benefits that go beyond their nutrient components. Therefore, nutrient database developers are being asked to provide databases that allow estimates of food group intakes as well as nutrient intake. Methodology for a standardized food-grouping system was recently developed by the United States Department of Agriculture (USDA). The Pyramid Servings Database (PSD) disaggregates most foods into their ingredients and then allocates each of these ingredients to one of 30 food groups. The Cancer Research Center of Hawaii has a food composition database with many unique foods because of our multi-ethnic population. We recently completed a project to add Pyramid Servings to each of the 1286 ingredients on our food composition database. The methodology we describe here may be useful to other researchers who want to calculate food group intake in a standard manner using the PSD but who must first link their database to the one from USDA. We highlight several foods on our database that were not present in the PSD or that did not match well to foods on the PSD. With a careful examination of energy and fat composition data, and the use of an energy adjustment factor, we were able to assign the appropriate number of Pyramid Servings to all foods.

© 2003 Elsevier Science Ltd. All rights reserved.

Keywords: Food composition tables; Pyramid servings; Multiethnic; Food groups; Dietary assessment

1. Introduction

Food composition tables traditionally include a variety of nutrients and other components of foods thought to be related to health. However, there is a growing awareness that non-nutrient substances in foods, in addition to their nutrient components, may have health benefits. For

*Corresponding author. Tel.: +1-808-564-5845; fax: +1-808-586-2982.

E-mail address: gsharma@crch.hawaii.edu (S. Sharma).

example, evidence suggests that soy products, rich in isoflavones, are associated with better plasma lipid profiles and may be protective against cardiovascular disease (Ho et al., 2000; Scheiber, Liu, Subbiah, Rebar, & Setchell, 2001). Therefore, nutrient database developers are being asked to provide databases that allow estimates of food group intakes as well as nutrient intake. In general, the endpoint of population-based nutrition research is to make recommendations for improving the diets of the target population. Thus, there are distinct advantages to looking at overall food consumption practices and translating recommendations into nutrition education messages focusing on overall diet (Kant, 1996).

Although it might seem straightforward to assign foods on a food composition table to appropriate food groups, several challenges exist. Perhaps the most important is the selection of a food group scheme. Many have been proposed in the past often based on differing food guides, which makes comparisons of findings across studies difficult (see Painter, Rah, & Lee, 2002, for examples of food groups used in various food guides).

Other general difficulties arise when attempting to describe food group intake. One is whether to express food amounts as grams or as servings. A limitation to using food weights alone is the difficulty of adjusting for differing moisture levels. For example, if unadjusted gram weights were used, then 100 g of dried apricots and 100 g of fresh apricots would contribute equally to fruit intake. However, because dried apricots have five times the dry solids of fresh apricots, it would be more logical to use servings so that the equivalent weight of dry apricots contribute five times as many servings as fresh ones. Another limitation of accumulating weights of foods is the difficulty in translating associations with disease outcomes into units that consumers understand. An additional advantage is that intakes can be directly evaluated by comparing servings consumed with servings recommended by educational tools such as the United States Department of Agriculture (USDA) Food Guide Pyramid (FGP) (USDA, 1992). For these reasons, expressing food intakes as *servings* per day, in addition to, or in place of grams per day has been appealing. This servings based approach may provide additional information to that obtained from accumulating food weights, and thus is a further method of determining diet and disease associations.

A related challenge is to decide how much of each food then counts as a serving. For example, how many grams of tomatoes should be considered a serving of vegetables? Another issue is how to handle foods that are mixtures. For example, pizza is composed of several ingredients from different food groups: crust (bread), cheese (dairy), tomatoes (vegetables), and sometimes meats as well. Thus, it is necessary to know the relative proportions of each ingredient so they may be assigned to the correct food group.

To address these issues, methodology for a standardized food-grouping system using 30 food groups was recently developed by the USDA. The Pyramid Servings Database (PSD) (released on line from USDA's Community Nutrition Research Group on the World Wide Web: <http://www.barc.usda.gov/bhnrc/cnrg>) disaggregates most foods into their ingredients and then assigns the appropriate number of FGP servings to each ingredient (USDA, 2000).

Nutrient database developers can take advantage of the PSD but must first match their foods with those on the PSD. The PSD provides servings data for foods reported by a representative sample of the US population. However, food composition tables for other countries or other populations may contain foods not on the PSD, and thus will require additional work to determine food group servings. The Cancer Research Center of Hawaii (CRCH) has a food

composition database with many unique foods because of our multi-ethnic population. We recently completed a project to add the Pyramid Servings to each of the 1286 ingredients on our food composition database. Here we present the methodology we used for this project. The goal was to calculate servings consumed for a large multi-ethnic cohort of over 215,000 participants that included African-Americans, Japanese-Americans, Hispanics, Native Hawaiians and Caucasians. Extending the PSD for this population sample is particularly challenging because many of the traditional foods consumed had not been reported in the US national surveys and therefore could not be directly matched to foods on the PSD. The methodology we describe may be useful to other researchers who want to calculate food group intake in a standard manner using the PSD but who must first link the USDA's PSD to their own food composition table.

The Food Guide Pyramid (FGP) is a guide designed to help Americans make more healthy food choices and reduce the risk of chronic disease (USDA & USDHHS, 2000; USDA, 1992). The FGP recommends a number of daily servings based on age and caloric intake for five major food groups (grain, fruit, vegetable, dairy and meat) and provides guidance on the avoidance of discretionary fat, added sugars and alcohol (Pyramid tip). The FGP was based on food choices to ensure nutrient recommendations are met. The PSD was developed to facilitate the examination of dietary data from national surveys and to allow comparisons of foods reported in those surveys with those recommended. The PSD allows direct comparisons between actual food group intake and the dietary recommendations of food groups in the FGP.

2. Methods

The CRCH food composition table (FCT) was developed as a research tool for nutritional studies in Hawaii. The nutrient data are drawn primarily from USDA databases, but many other sources are used as well. Other sources were needed because USDA's nutrient databases did not contain many of the foods consumed in Hawaii; because national surveys are not conducted in Hawaii, local foods were often not included. Furthermore, the coding scheme used for the CRCH FCT was not the same as the PSD coding scheme, which is based on the USDA's Food Coding Database for the Continuing Survey of Food Intakes by Individuals (CSFII). Thus, to take advantage of USDA's PSD, we needed to link the PSD to the CRCH FCT (Fig. 1).

2.1. Pyramid servings database

The PSD file (USDA, 2000) identifies the number of servings from 30 food groups provided per 100 g of a food item for over 7000 foods reported during the 1994–1996 CSFII (USDA, 2000). Twenty-two of these food groups are subsets of the five main food groups in the FGP (grains, vegetables, fruits, dairy and meat), and a further three groups make up the Pyramid Tip (added sugar, discretionary fat and alcohol) (see Table 1). Serving sizes were determined using the guidelines from the FGP (USDA, 1992) because these are familiar to most consumers (e.g., a slice of bread, a medium fruit). Unlike the servings used in the US for labeling purposes (Kurtzweil, 1993), the FGP servings are not an average amount that is consumed at a meal or snack. Further details on serving sizes for specific foods are available (USDA, 2000).

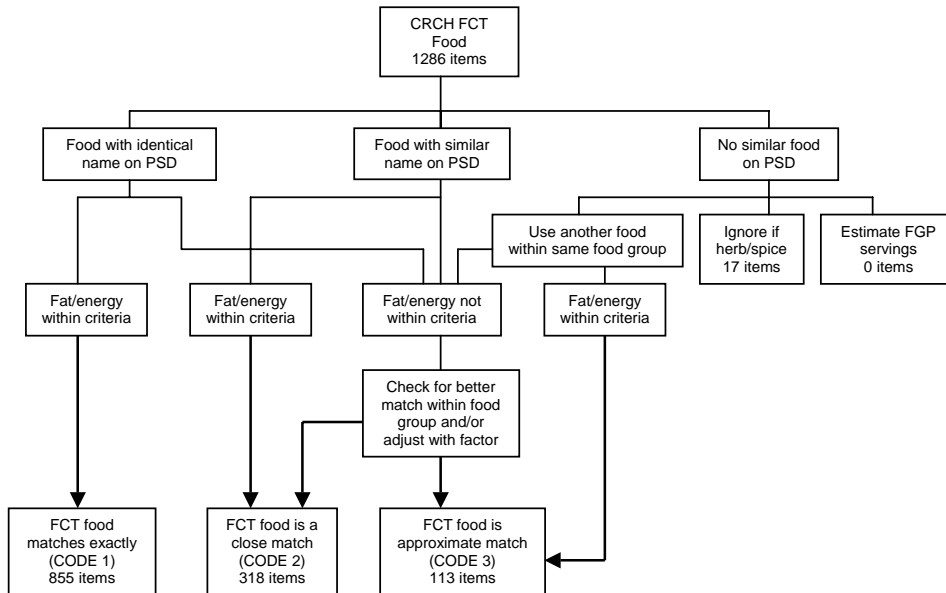


Fig. 1. Procedure for assigning Food Guide Pyramid Servings to the CRCH FCT foods.

Unlike the main food groups, the three groups in the Tip are not measured in servings per 100 g. Added sugar is given in teaspoons per 100 g of food (one teaspoon being 4 g), discretionary fat is in grams per 100 g of food, and alcohol is in drinks per 100 g (one drink as defined by the Pyramid contains 13–14 g of alcohol). Added sugars include all sugars eaten separately or added at the table as well as sugars that are used in food preparation or processing. ‘Discretionary fat’ is defined as excess fat from the five major food groups that would have been avoided if only the lowest fat forms of the foods had been selected, as well as any additional fat added to foods, such as butter or salad dressing.

2.2. Cancer Research Center of Hawaii Food Composition Table

The CRCH has a unique and extensive food composition table containing foods commonly eaten by ethnic groups in both Hawaii and California; it is regularly updated and expanded. Currently the database contains over 1500 food ingredients of which 1286 were used to quantify intakes of the multi-ethnic cohort participants. A recipe database provides information on ingredients and proportions for an additional 700 mixed foods. Prior to calculating intakes, foods that are mixtures are disaggregated into their ingredients using the recipe database.

2.3. Methodology for matching the USDA’s survey database food codes to the Cancer Research Center of Hawaii’s Food Composition Table

The goal of this study was to find an appropriate match on the PSD for each of the 1286 food ingredients reported by our study participants. Because the food codes used for the CRCH FCT

Table 1
Pyramid Food Groups

Major food groups	Pyramid food groups
Grain group	1 Total grain
	2 Whole grain
	3 Non-whole grain
Vegetable group	4 Total vegetables
	5 Dark green leafy vegetables
	6 Deep-yellow vegetables
	7 White potatoes
	8 Other starchy vegetables
	9 Tomatoes
	10 Other vegetables
Fruit group	11 Total fruits
	12 Citrus fruits, melons, and berries
	13 Other fruits
Dairy group	14 Total dairy
	15 Milk
	16 Yogurt
	17 Cheese
Meat and beans group	18 Meat, poultry, fish
	19 Meat (beef, pork, veal, lamb, game)
	20 Organ meats (meat, poultry)
	21 Frankfurters, sausage, luncheon meats (made from meat or poultry)
	22 Poultry (chicken, turkey, other)
	23 Fish (fish, shellfish, other)
	24 Eggs
	25 Cooked dry beans and peas ^a
	26 Soybean products (tofu, meat analogs)
27 Nuts and seeds	
Pyramid tip	28 Discretionary fat
	29 Added sugars
	30 Alcohol

^aCan be counted toward the meat group or the vegetable group, according to the Food Guide Pyramid.

are not the same as those used for the PSD, initially the matching was based on the food name. A further check was carried out to determine if any of the matched foods differed by more than 20% in energy content (calories) per 100 g. Foods that did not match to within 20% were looked at in more detail to determine if a better match could be found, based on energy. As a further check on the quality of matching, the fat content per 100 g for all matched foods was compared and the match was investigated if the difference was greater than 20%. Even if energy and fat were similar we did not match across food groups, e.g., a vegetable with a fruit, to preserve the intended food-grouping scheme.

Matching on fat content was important because any discrepancy would contribute to errors in the discretionary fat variable in the Pyramid Tip. For example, because non-fat milk is available in the US, milk that contains any fat (reduced fat and whole milk) has all fat allocated to the discretionary fat variable. A further example is fat from meat, USDA chose 9.35 g of fat per 100 g of food item, as the lowest fat form of meat available, and any fat in excess of this contributes to the discretionary fat variable. If a CRCH FCT food contained less than 5 g of fat per 100 g, we considered a match with a Pyramid Servings database food that was also low in fat appropriate, as long as the items belonged to the same food group.

Although added sugar is also a variable in the Pyramid Tip and could be used for matching foods, we do not carry this variable in the CRCH FCT. We chose not to match micronutrient content because the ultimate goal was to calculate food group intakes, not micronutrient intake.

The PSD contains mostly foods in the ‘as consumed form’ and therefore does not contain some uncooked foods (like dried egg) which are on the CRCH FCT. In order to match these PSD foods to the uncooked or dehydrated form in the CRCH FCT, an energy factor was calculated and used to adjust the number of Pyramid Servings. The energy factor thus allowed a more accurate calculation of the Pyramid Serving per 100 g of a CRCH FCT food.

A code system was developed to denote the quality of the match between the food item on the PSD and the CRCH FCT: (1) an exact match, e.g., the same food item was present in both databases; (2) a similar but not identical food was found, and details of the slight differences were documented for each food under a comments column; (3) a match was found within the same food group but the foods were not similar and there may be differences in the micronutrient content. Differences in micronutrient content were not important for this particular study, but we wished to identify these dissimilar matches for use in future work to obtain nutrient data from the USDA’s survey nutrient database.

3. Results

We were able to find a match in the PSD for most foods using the food name and nutrient composition. The code system developed to denote the quality of the match between the PSD and the CRCH FCT identified 855 items with code 1 (an exact match), 318 items with code 2 (a similar but not identical food was found), and 113 items with code 3 (a match was found within the same food group but there may be differences in the nutrient content). Even for foods given a code 3, the matches were considered acceptable. Thus we did not need to estimate PS for any foods.

Table 2, item 1, gives an example of a food (casaba melon) present in each database that matched exactly and the match code was set to ‘1’. However, for some foods, such as ‘tofu soybean curd’, although both databases contained the same name food, the nutrient composition was different (Table 2, item 2). In this case, although the differences in percents were large, the absolute differences were relatively small and we considered the match acceptable but the match code was set to ‘2’. For other foods the food names were not identical but the nutrient composition matched exactly. For example, ‘soy nuts’ in the PSD database exactly matched ‘soy bean, roasted, salted’ in the CRCH FCT (Table 2, item 3).

The foods that were more challenging in terms of finding the best match, but still keeping within the same food group tended to be the traditional Asian foods. National food surveys in the US

Table 2
A comparison of selected CRCH foods with the best match from USDA's PSD

#	PSD description	CRCH description	PSD (kcal/100 g)	CRCH (kcal/100 g)	% difference (kcal)	PSD fat (/100 g)	CRCH fat (/100 g)	% difference fat	Energy factor ^a	Match code ^b
1	Melon, casaba	Melon, casaba	26	26	0	0.1	0.1	0	1	1
2	Soy bean curd	Tofu soy bean curd	76	59	29	4.8	3.0	57	1	2
3	Soy nuts	Soy bean, roasted salted	471	471	0	25.4	25.4	0	1	1
4	Squid raw	Bagoong sauce (small fish)	92	116	21	1.4	1.8	23	1	3
5	Tuna fish smoked	Eel raw	201	184	9	10.1	11.7	13	1	3
6	Chinese gravy	Pork blood, cooked	42	55	24	0.3	0.2	22	1	3
7	Soy nuts	Tofu, freeze dried	471	480	2	25.4	30.3	16	1	2
8	Soy bean curd	Okara (tofu residue)	76	77	1	4.8	1.7	176	1	2
9	Soy beans, cooked, no fat	Tofu gan pressed	172	169	2	8.9	9.5	6	1	2
10	Kale, cooked, no fat	Bittermelon leaves (paria)	28	30	7	0.4	0.7	42	1	3
11	Fresh egg	Dried egg	149	594	75	10	41	76	4	2

PSD = Pyramid Servings Database; CRCH = Cancer Research Center of Hawaii.

^aEnergy factor is used to adjust Pyramid Servings to match the CRCH Food Composition Table food.

^bMatch code 1 = exact match; 2 = similar but no identical food was found (details are given in the comments column of the database); 3 = match was found within the same group but there may be differences in nutrient content (details are given in the comments column of the database).

target African Americans and Hispanics but not specifically Asian groups. The PSD was developed based on national survey data and many foods consumed by Asians were therefore not included. For example, an Asian food on the CRCH FCT, 'bagoong sauce' (a fish sauce), was matched to 'squid, raw' on the PSD, although the difference in the energy match was 21% (Table 2, item 4). Further examples of such foods that were not included in the PSD include 'raw eel', which was matched to 'smoked tuna fish', and 'cooked pork blood', which was matched to 'Chinese gravy/stock' (Table 2, items 5 and 6). Note that all three of these matches were assigned a match code of '3', indicating the micronutrient content was likely to be different.

Soy and soy products were frequently consumed in many forms by the Japanese American and Native Hawaiian populations in Hawaii. However, the PSD contained relatively few soy items, and matching foods was therefore more challenging. For some foods, such as freeze-dried tofu, the difference in energy and fat was within 20% when matched to 'soy nuts' in the PSD (Table 2, item 7) so we considered the match acceptable without using an energy adjustment factor. However, for other soy products the difference was much larger. For example, if 'soy bean curd' in the PSD was matched to 'plain tofu residue' on the CRCH FCT, the difference in energy was only 1.0% but the difference in fat was 176% (Table 2, item 8). However, because the absolute difference was only 3.1 g of fat per 100 g of food, again we considered the match acceptable. Pressed tofu gan (Table 2, item 9) illustrates a case where a good match was found for both energy (within 2%) and fat (within 6%). Because the names of these foods did not match exactly, they were assigned a match code of '2'.

Examples of foods that differed by more than 20% in terms of energy or fat content, were vegetables where the fat and energy content was low. For example, 'kale' in the PSD was matched to 'bittermelon leaves' even though the difference in fat between the PSD food and the CRCH FCT was 42%, because the absolute difference was only 0.3 g of fat (0.7 g vs. 0.4 per 100 g) (Table 2, item 10).

The CRCH FCT contained 10 dried foods that required an energy adjustment factor because they were listed as the fresh form on the PSD. For example, the CRCH FCT contains dried egg, whereas the PSD lists only fresh or frozen raw egg (Table 2, item 11). Dried egg contains 41 g of fat/100 g and 594 kcal/100 g while fresh eggs contain 10 g of fat and 149 kcal/100 g. To match the PSD fresh egg to the CRCH FCT dried egg, a factor of 4.0 was used to 'concentrate' the number of Pyramid Servings per 100 g. In this way the correct number of Pyramid Servings was allocated.

Further examples of foods that we found difficult to match were seaweeds and fishes, because there were many different forms and varieties on the CRCH FCT, but far fewer on the PSD. However, in all cases we were able to find acceptable matches using fish or seaweed from the PSD.

Seventeen items on the CRCH FCT were herbs and spices that were not in the PSD database. We decided not to assign these items to food groups because almost all are consumed in very small amounts and contain few nutrients. These foods included: baking soda, ground ginger, peppers, pimento, salt, cloves, nutmeg, oregano, allspice, cinnamon, basil, chili powder, curry powder, monosodium glutamate, paprika, black pepper and pepper cayenne.

After all 1286 foods were matched, a linkage database was created with the PSD food code, CRCH FCT food code, the energy factor, and the quality of match code. This database was then used to compile the 30 Pyramid Servings variables for each CRCH FCT food ingredient. The CRCH FCT is now used by a variety of studies to calculate intakes of Pyramid Servings.

4. Discussion

We have described how the PSD can be linked to other food composition tables and used in different populations. Foods must be carefully matched so they are within the same food group. In addition differences in energy content must be considered so that the number of servings is correct, and differences in fat and sugar must be considered so discretionary fat and added sugar are correct. We have highlighted several foods that are not present in the PSD and made recommendations on how such foods, if consumed in other surveys, can be linked to the PSD, so that a similar number of Pyramid Servings can be allocated for those foods. Daily Pyramid Serving intakes could thus be calculated in addition to daily nutrient intakes when analyzing dietary data. We have focused on the methodology for extending our particular CRCH FCT, but a similar method could be used for other FCTs.

The PSD is a unique database and provides an opportunity to present data from different populations using a similar food-grouping system. In this way, results from dietary surveys will be more comparable. If associations are determined between food groups and disease outcomes, conclusions drawn will be more meaningful than at present, where different research groups utilize many different systems of grouping foods.

The PSD can be used by populations from countries other than the US to provide a standard method of grouping foods, even though these countries may have their own nutrition-related problems, food composition tables and dietary recommendations. While comparisons with the US FGP recommendations may not be appropriate, reporting food intake using a standardized grouping methodology would enable direct comparisons of findings across countries, thus leading to more conclusive associations between dietary patterns and disease outcomes. However, there are clearly some challenges in using the PSD in other countries. Equivalent foods may not be available for many country-specific foods (e.g., if the meat is of a lower or higher fat content than that available in the US, or if the amount of added sugar in processing varies). A further limitation is that the composition of foods in the US may be very different from those elsewhere, even though the names may be the same. For example, the ingredients for pizza reflect the composition of pizza in the US, based on national surveys over several years. Pizza in other countries may be quite different.

Even diets that differ to a large extent from the US diet could use the PSD approach for comparison of the broader food groups. For example, while the foods within each food group may be different (e.g., the foods within the grain group), the total grains would be comparable, as would other broad groups like vegetable, fruit, dairy or meat (including meat alternatives such as dried beans and tofu). Virtually all foods can be assigned to these groups. In theory, food group servings could be estimated for additional foods, providing USDA's criteria were followed. Consideration would need to be given to food group assignments (e.g., does a new vegetable go into a defined vegetable subgroup or into the 'other vegetable' category?) and also to the components of the food that go into the Pyramid Tip (e.g., exactly how many grams of fat in a food are allocated to discretionary fat or what kind of sugar would be allocated to the 'added sugar' category?). Because we were able to find acceptable PSD matches for all the CRCH FCT foods that were reported by our study participants, we did not need to estimate Pyramid Servings for any of our foods.

Researchers could also develop new food grouping approaches that more closely match local food guides. The PSD approach could serve as a model for allocating each food on the local food

composition table into these new food groups. However, this is a very time consuming and demanding process and the results would not be comparable to those reported by other studies. The USDA PSD was developed over many years and uses very detailed, well documented criteria based on diets and food availability in the US.

We are currently using this database in many ways, firstly to assess adherence to the FGP recommendations in each ethnic group in our multi-ethnic cohort and secondly, to assess associations between diet and cancer. We are determining if the degree of adherence to the FGP is associated with risk of cancer and which food groups may be the most protective, as well as how many servings provide the most protection.

5. Conclusion

The PSD can be adapted and used to assess the diets of ethnically diverse populations. Some effort may be required to match some local foods to foods on the PSD but we were able to find acceptable matches for all 1286 of our foods. We are now able to compare intakes of our study participants with the FGP dietary recommendations. These data will also be useful for further developing dietary guidance for ethnically diverse populations and for determining food-level diet–disease associations in these groups. The PSD is available to all researchers from the USDA's Community Nutrition Research Group. We believe other research groups will also find many advantages to developing databases that allow computation of intakes using these standardized FGP servings variables.

Acknowledgements

We are most grateful to Mary Arakaki for her help with matching the foods in the two databases.

This research could not have been undertaken without the financial support of the National Cancer Institute (grant number NCI R01 CA54821).

References

- Ho, S. C., Woo, J. L., Leung, S. S., Sham, A. L., Lam, T., & Janus, E. (2000). Intake of soy products is associated with better plasma lipid profiles in the Hong Kong Chinese population. *American Society for Nutritional Sciences. Journal of Nutrition*, 130, 2590–2593.
- Kant, A. K. (1996). Indexes of overall diet quality: A review. *Journal of the American Dietetic Association*, 96, 785–791.
- Kurtzweil, P. (1993). Nutrition facts to help consumers eat smart. US Food and Drug Administration FDA consumer. May: 22–27.
- Painter, J., Rah, J., & Lee, Y. (2002). Comparison of international food guide pictorial representations. *Journal of the American Dietetic Association*, 102, 483–489.
- Scheiber, M., Liu, J., Subbiah, M., Rebar, R., & Setchell, K. (2001). Dietary inclusion of whole soy foods results in significant reductions in clinical risk factors for osteoporosis and cardiovascular disease in normal postmenopausal women. *Menopause*, 8, 384–392.
- United States Department of Agriculture (1992). The Food Guide Pyramid. Home and Garden Bulletin No 252.

- United States Department of Agriculture (2000). US Department of Agriculture Agricultural Research Service. 1994–96 Continuing Survey of Food Intakes by Individuals and related survey materials. Pyramid Servings. Food Surveys Research Group. CD ROM USDA, Riverdale, MD 20737.
- United States Department of Agriculture and United States Department of Health and Human Services (2000). Dietary Guidelines for Americans. Fifth Edition. Home and Garden Bulletin No. 232. USDA, Washington, DC.