

## Adaptation of international nutrition databases and data-entry system tools to a specific population

Iris Shai<sup>1,2,3,\*</sup>, Hillel Vardi<sup>2</sup>, Danit R Shahar<sup>1,2</sup>, Ayelet B Azrad<sup>1,2</sup> and Drora Fraser<sup>1,2</sup>

<sup>1</sup>S Daniel Abraham International Center for Health and Nutrition, Ben-Gurion University of the Negev, PO Box 653, Beer-Sheva 84105, Israel; <sup>2</sup>Epidemiology and Health Services Evaluation Department, Faculty of Health Sciences, Ben-Gurion University of the Negev, Beer-Sheva, Israel; <sup>3</sup>The Nutrition Department, Harvard School of Public Health, MA, USA

Submitted 13 June 2002; Accepted 14 November 2002

### Abstract

*Objective:* To develop a nutritional dietary intake database based on available reliable international nutritional databases adapted to the local needs of a specific population.

*Design:* The Negev Nutritional Study (NNS) is a survey of a random sample of the Negev population regarding their dietary intake using 24-hour dietary recalls. A nutritional database for the Israeli population was developed based on adaptation and modification of the US Department of Agriculture's database. A data-entry system was developed based on the logic of the US Food Information Analysis System. The system was designed as bilingual (English and Hebrew). Local foods and recipes were collected during the NNS, which included 1465 24-hour diet interviews.

*Results:* During the course of the NNS, 383 basic Israeli recipes were constructed. In total 1362 Israeli products were added to the database, and each was given a code, specific gravity and portion size. Most of the added products were cereals and grains and dairy products. The added recipes were collected from the interviewees in the NNS and from the most popular cookbooks.

*Conclusions:* This paper describes the process undertaken to develop an Israeli food composition database as well as the data-entry system. This knowledge may aid other research groups in developing a computerised, nation-specific nutritional database and data-entry system adapted to their own specific local needs.

**Keywords**  
Nutrition database  
Nutrition database adaptation  
Database data-entry tools

Despite the general trend towards globalisation, it seems in the near future that nations and cultures will continue to preserve their own nutrition customs. People from different populations will continue to eat food prepared using traditional recipes and locally manufactured foods will remain different due to local regulations regarding food fortification, food additives, etc. Therefore nutritional databases cannot be applied globally unless a methodical local adaptation is performed.

In Israel, with its varied population, it was necessary to develop a unique nutritional database that reflected the rich ethnic texture of the country. The most suitable way was to base it on population survey data. Israeli foods, both those prepared at home and ready-made foods bought at supermarkets, differ from American or European foods. For example, the Israeli market has a varied selection of low-fat dairy products and cheeses that are not available elsewhere and thus are not described in Western databases.

The local sources for developing a nutritional database were limited. The Department of Nutrition of the Ministry

of Health published a book of nutritional component tables in 1996<sup>1</sup>. The values in the tables were primarily based on the United States Department of Agriculture (USDA) tables<sup>2</sup>, which meant that ethnic recipes and local brand name products were not included. No national nutritional database has ever been compiled in Israel. A project to develop such a database is currently ongoing in the Nutrition Department of the Ministry of Health, but has not yet been published. In the absence of a national database and within the framework of establishing and validating nutritional research tools for the Israeli population, the need arose to develop an Israeli food database. The decision was to base the dataset on available international databases adapted and modified for the local ethnic population.

The SR12 version nutrition database was developed from the United States Nutrition Data Bank (USNDB) by the USDA<sup>3</sup>. It was chosen as the preferred starting point for several reasons. First, it is regarded as one of the most extensive and reliable databases and is updated regularly. Second, many foods are imported from the USA to the

\*Corresponding author. Email irish@bgumail.bgu.ac.il or ishai@hsph.harvard.edu

Israeli market. Third, SR12 is available online to the public (<http://www.nal.usda.gov/fnic/foodcomp/Data/index.html>). Finally, the methods used in developing and updating the USDA database have been fully described<sup>4-6</sup>.

A nutritional database consists of several tables<sup>7</sup>. One is for food codes and can include descriptions of foods, food brands, descriptions of measurement and coefficients for converting volume measurements to weights. Another is a table of nutritional components that, in addition to values of the components, may include retention factors for estimating changes during cooking. A recipe table includes the recipes constructed from primary products for the purpose of calculating the content of food components, which could include information on changes in moisture and fat content during cooking.

The aim of this work was to adapt reliable available nutritional databases in order to allow analysis of dietary intake for use with the local language (Hebrew), tailored to the needs of the native population. Our experience may help other research groups that are developing a computerised, nation-specific nutritional database and data-entry system adapted to their own specific local needs.

## Methods

### *Developing a data-entry system*

The data-entry system developed (using Microsoft Access<sup>®</sup>) was based on the logic of the American Food Information Analysis System (FIAS) program<sup>8</sup>. Our system was designed as bilingual (English and Hebrew) for both tables and interface. The system guides the user in the process of entering data. Initially, the coder enters demographic data for the subject. Then, for each food consumed, information regarding the food and the occasion is added. The system displays items chosen using parts of the food name, thus allowing it to match the reported food in the questionnaire. In addition, the coder can check the nutrient values, the available weights or the recipe ingredient of the selected food item. Recipes can be modified on-line. The coder can obtain the total nutrients and the weight calculation report at any stage of data entry. The Israeli data-entry system can be updated with new commercial products, can be used to construct recipes, to make specific modifications to existing recipes and to calculate weights from descriptions of volume. In addition, it allows the use of alternative databases and to export data for statistical analysis. The system can present the data on food nutrients for each food item (as they are reported separately or as part of a recipe) as well as the total nutritional intake of the interviewee.

### *Establishing the nutritional database*

Establishing the database consisted of three stages: the first was modifying existing datasets for local use; the second was collecting information on Israeli commercial food

products; and the third was adding Israeli recipes to the database.

### *Adapting the USDA food composition database for local use*

The database tables used were from SR12 and The Continuing Survey of Food Intakes by Individuals (CSFII). We retained the following: (1) food names and codes, (2) nutrient value per 100 g, (3) food weights, (4) recipe ingredients and (5) retention factors for nutrients during cooking, as well as fat and moisture loss/gain during that process. The names of the foods from the USDA were translated into Hebrew using a translation program and then each item was individually checked. Unfortified raw food items (e.g. flour, rice, nectars) were identified for future use. Shortening as a recipe ingredient was replaced by margarine or soy oil according to the nature of the product.

### *Collecting information on Israeli commercial food products*

This process has been ongoing since 1998. Initially, a market survey was prepared and then products were added as required for analysis. In the first Negev Nutrition Study (NNS), which included 1465 adult 24-hour recall questionnaires, we used data from food manufacturers and professional and commercial websites, in addition to using food labels. A limited number of products tested at laboratories were received from the Nutrition Department of the Ministry of Health.

Portion size<sup>9</sup> determinations were performed in one of two ways. Weight data, including package weight, data on portion units in the package and the weight of each portion unit were collected from the food labels. Alternatively, portion size was calculated by dividing total package weight by the number of portions. Foods consumed in units (like cookies, sausages, bread slices, etc.) were weighed three times using a digital scale before calculating the average weight per unit. For some of the Israeli products unit size was measured (e.g. cookies, hard cheeses and sausages). The aim of describing size was to build weight descriptions of the general product whose brand name is not specified, for example a 3 × 5 cm cookie. The specific gravity of each Israeli product was obtained from the USDA, etc. Direct measurements were performed for selected food items (e.g. ice cream, cereals) by weighing some of the foods in a container of known weight. Measurements with a ruler were also carried out at the homes of the interviewees where appropriate.

A code was given to each product. The 8-digit code number was based on a code for the food data group, cooking method, fat or calorie enrichment or reduction, manufacturer's name, and following the USDA method<sup>10</sup>.

Completing (imputing) the nutrient data for Israeli food products<sup>11</sup> was the last and the most complex stage.

Foods were individually matched to a similar food from the USDA<sup>2</sup>, British<sup>12</sup> or Italian<sup>13</sup> data. The degree of similarity between products was checked by looking at minimum and maximum values of differences in contents of carbohydrates and protein between Israeli products and the base products. In analysing different types of local white spreadable cheese, for example, we found a consistent association of 3 mg of cholesterol for each 1% of fat. Calculation of fat composition (saturated, monounsaturated and polyunsaturated) was carried out according to the ratio in the base product, adjusting for the percentage of fat in the Israeli product. An alternative method was to recalculate the recipe that best represented the commercial food. For example, for 1.5% fat yoghurt with chocolate-covered cereal, a recipe was constructed as follows: 1.5% yoghurt, cereal and chocolate in amounts that express the content of the reported macronutrients. The final decision was made only after the suggested modification was approved over the course of several sessions. The potential base products were chosen from the databases by a staff of dietitians. In the second stage a broader forum of dietitians discussed the various possibilities. This forum determined the principles of modification and decided which base food was most suitable. In the final stage, the research staff approved the modifications.

#### *Building a database of Israeli recipes*

Recipes were collected from 14 popular Israeli cookbooks which tended to recommend popular recipes or recipes typical of certain ethnic groups. From this source, 130 basic recipes were chosen such as *falafel* (deep-fried ground chick peas), doughnuts, meat or vegetarian *bamin* (Sabbath casserole), chicken soup, gefilte fish (ground fish patties) and so on. Additional recipes from 24-hour recall interviews were added or existing recipes modified. In this way, the new immigrants' recipes, e.g. Ethiopian and Russian, were recorded. From a number of descriptions of the same food in the reported 24-hour recall questionnaire, an 'average' recipe was chosen and 'cooked' in the program. Modified recipes were saved in a separate modification file. Modification decisions were made at the supervisory level, similar to the procedure at the USDA<sup>5</sup>.

Recipes were constructed as follows.

1. *A base recipe was chosen.* The aim was to choose a base recipe that included the main ingredients in the Israeli recipe and used a similar cooking method, thus ensuring accurate calculation of percentages of moisture loss and fat absorption. For example, in order to add a recipe for vegetarian *bamin* based on beans, a long-cooking Mexican casserole recipe containing beans was chosen as the food base. If, after a thorough search, no suitable base recipe was found, the percentages of moisture loss and fat absorption were chosen from the most similar cooking method.
2. *Coding.* The recipes were coded using a 6-digit code number<sup>10</sup>. In cases where a number of Israeli recipes were developed based on one base recipe, the last digit of the code was changed according to sequence numbers.
3. *Naming recipes.* Each recipe was given a title that included its description and source.
4. *Choosing the recipe's ingredients.* The base recipe was modified to produce an Israeli recipe. The ingredients were carefully chosen from the 3000 raw foods that were described in the SR12. The guiding principles were to choose unenriched foods, as they exist in Israel, and raw foods.
5. *Determining the ingredient weights.* The guiding principle was to choose foods in their original volume, before processing; for example, two cups of uncooked dry white rice, according to the description of quantities in the recipe. The program was used to convert volume to weight in grams.
6. *Calculating the retention factor.* For each ingredient in the recipe, the quantity of vitamins and minerals lost during cooking was calculated according to the cooking method described (e.g. boiling, steaming, baking, heating at a certain temperature, etc.).
7. *Calculating the percentages of moisture loss and fat absorption.* Percentages were determined by the program, based on the cooking method described.
8. *Determining the portion size.* The total weight of a recipe (after allowing for moisture loss and fat absorption) was divided by the number of portions described (e.g. 40 cookies per recipe). The calculated standard portion was verified by weighing samples of the food and by comparison with the standard portion size<sup>14</sup>.
9. *Determining weight/volume ratio.* At this last stage, specific gravity was determined for each recipe with the aim of converting volume models from the quantity guide used in the 24-hour recall interview to units of weight. The specific gravity was chosen using a recipe of similar texture or direct volume and weight measurements.

At the end of the process, the nutritional value of the chosen recipes was calculated, taking into account all factors described.

#### **Quality assurance**

The quality assurance of the newly adapted database was conducted at three levels: (1) interviewer and coder; (2) database; and (3) calculation of nutrient values by the program.

#### *Interviewer and coder*

Each completed questionnaire was reviewed by a member of the data management team, distinct from the interviewer. The questionnaires were edited and checked

for missing data within one to three days following the interview.

Quality control of data entries was performed in three cycles. The first and second cycles were repeat checks by dietitians who re-examined the interview data per questionnaire, the food items chosen and their portion size. The third cycle involved surveying the entire database to detect unusual nutrient or food model values.

#### Database

Nutritional values were compared with the work done separately by the Israeli Ministry of Health, Nutrition Department. The only discrepancy was shown in calcium values since some of the American foods used in our database were fortified with calcium. Values were corrected according to the values of the Ministry of Health.

#### Nutrient calculation

The new Israeli recipes were entered into the computer twice, once using the USDA program FIAS<sup>8</sup> and then using our new system. Nutritional values were compared and no difference was detected between the two programs.

Dietary intakes of a sample of people participating in the study who did not consume Israeli foods were entered in the newly developed program and re-entered using FIAS. The same nutritional values were calculated using both programs.

#### Results

Table 1 presents the available bilingual food databases. During the project, more than 380 basic Israeli recipes were constructed and data on more than 1360 Israeli commercial products were collected. During the NNS, more than 1200 specific modifications were done to the basic recipes. Most of the added recipes were from the grain and cereal group, the vegetable group and the meat, poultry and fish group (Fig. 1). The main recipe groups were vegetable soups and salads, cookies and cakes, mixed meat and fish dishes, and mixed dough dishes. A quarter of the sources of data for recipes were popular recipes added from oral information, the remainder were taken from 14 popular cookbooks. The Israeli products that were added to the product

database were, for the most part, from the cereals and grains group, the milk and dairy products group, the vegetable group and the meat/poultry/fish group (Fig. 2). In the cereals and grains group, bread (12.1% of all products), cakes (7.5%) and crackers (7.4%) were particularly prominent. In the milk and dairy product group, ice cream (14.1%), milk (7.9%) and cheeses (7.1%) were predominant.

#### Discussion

As a result of this study, we have a bilingual Israeli food composition database that allows for the analysis of 30 nutrients and 11 000 foods, including Israeli products and recipes. The system also allows modification of base recipes according to the report of each interviewee.

The basic needs of nutrition research are validated dietary assessment tools and an adapted nutritional database that fits the characteristics of the study population. To this end, data from a population survey are required. The 24-hour recall, an open questionnaire that collects information on the foods consumed by a subject on the previous day, is widely used for this purpose<sup>15,16</sup>. While local foods and recipes must be obtained in this way, many basic food items in countries that import some of their products, such as wheat, can be obtained elsewhere, thus saving time and effort.

The Israeli diet, like the Israeli marketplace, differs from that of North America and Europe. The differences stem, for example, from differences in iron, folic acid, calcium and vitamin C enrichment levels common in some Western food products. Another example is the higher content of fat in the cheeses and other dairy products used elsewhere. These differences required the development of the country-specific food composition table, a need not unique to Israel, but common to localities establishing dietary assessment methodologies.

The database that we developed was based, for the most part, on the USDA database. This database, the USNDB, is the most highly developed and accessible database in the world today. In building a local database, it is not always necessary to develop a new system of analysis, since today there is a variety of accessible databases available for reference in the *Nutrition Data Bank Directory*<sup>17</sup>.

**Table 1** Bilingual databases available on 1 April 2002

Source	Number of recipes	Process
USDA raw food items	3033	Translated, unfortified
USDA recipes	4854	Translated, unfortified
USDA one-line recipes	2397	Translated
Israeli commercial products	1362	Missing data completed
Israeli ethnic recipes	383	Computed from basic products
Israeli modifications performed on recipes obtained from 24-hour interviews in the NNS	1244	Modified from existing recipes

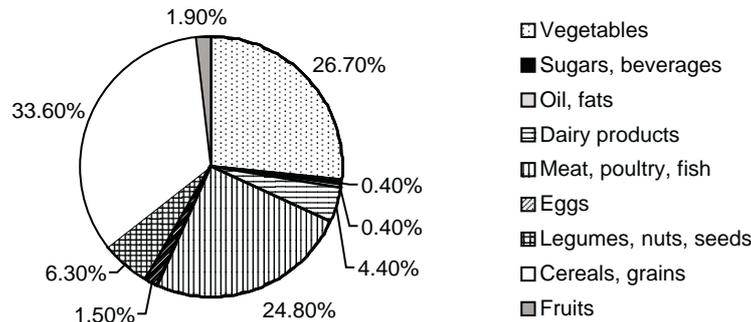


Fig. 1 Distribution of additional local recipes to the database (by food group),  $n > 380$

The American<sup>18</sup>, French<sup>19</sup> and Canadian<sup>20</sup> food tables are based on a number of sources. The Canadian food tables, called *Nutrient Value of Some Common Foods*<sup>20</sup>, are based on USDA pamphlet number 8, with specific adjustments<sup>21</sup>. At the USDA, a complex system of databases for processing and analysing national nutritional survey data in the USA is used. This system has been in use since the 1970s and was recently employed by the CSFII from 1994 to 1996<sup>22</sup>. Of the recipes we added, one-fifth of the ethnic recipes were derived from vegetable soups and salads. Commercial food products were, for the most part, from the cereals and grains groups (mainly breads, cakes and crackers) and from the dairy group (mainly cheeses, ice cream and dairy desserts). The need to add these foods is reflective of the major differences between the Israeli and the North American markets, which are, as previously noted, more enriched flours and fatter milk products in the USA than in Israel.

In this task of developing dietary assessment in a population that had not been thoroughly studied previously, the decision was made to use resources available today. The USDA food composition table was the basis of the only existing food table in Israel. In addition, many basic foodstuffs such as wheat are imported to Israel and not grown locally. Israeli supermarkets import a selection of US manufactured products such as breakfast cereals, while other US food companies have established manufacturing plants (e.g. Coca Cola) and restaurants (Pizza Hut, McDonalds, Burger King, etc.) in Israel. The

availability, comprehensiveness, careful compilation and testing by the USDA, and its previous use by the Ministry of Health, made the USDA food composition table the starting point of choice for our project.

One major limitation was that our source of data on the contents of food products was food product labels. In the absence of the option to carry out broadly based chemical tests and vitamin tests in particular, food labels also serve as a data source at the USDA. We should be aware of this drawback, even though it is not unique to our nutrition database. In any database, there are a number of sources of data. In England<sup>12</sup>, the composition of almost all foods was determined by tests at the British government's chemical laboratory. After completion of the English food tables in 1978, additional information about certain foods was published in professional journals and as an appendix to the tables, e.g. particles of dietary fibre (such as non-cellular polysaccharides or cellulose), vitamin C content, the composition of mother's milk, amino acids, fatty acids. In a number of countries (e.g. Cuba), the nutritional values of local foods were adjusted directly from nutritional knowledge on the basis of foreign food tables that had been checked and validated in the past<sup>23</sup>. In Ethiopia, Western Africa and Zambia, the values of a number of foods were determined by laboratory tests, although the components and the samples were few<sup>24</sup>. When there are no data on the value of certain nutritional components, chemical tests are carried out directly to determine their level. Such validation cannot indicate perfect adjustment due to differences in salt

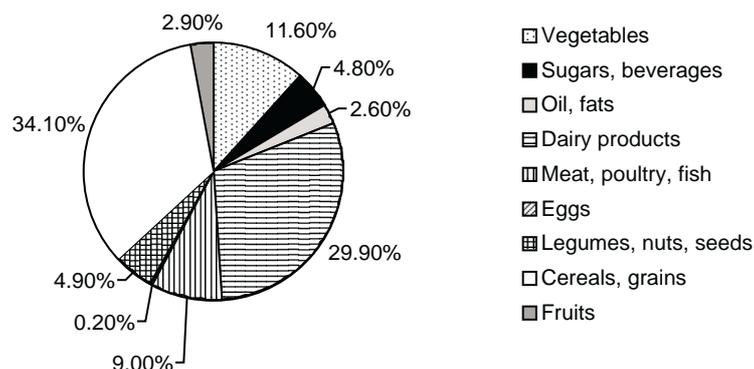


Fig. 2 Distribution of additional local products in the database (by food group),  $n > 1300$

content, processing of water-soluble vitamins, enrichment, soil composition, water and type of animal feed used. Food values derived from food tables and chemical tests represent the maximum values available to the body, but not necessarily the amount actually absorbed.

In an effort to prevent some of the potential errors in food tables, the International Network of Food Data Systems (INFOODS) was established<sup>24</sup> for the purpose of comparing and correcting databases around the world. This system encourages the development of food databases world-wide. Rand<sup>25</sup> describes the aims of this network. Nutrition database management, as it is managed in Europe<sup>26–28</sup>, requires careful examination and endless updating to reflect the changes that occur over time. Our future goal is to update our database, to add to it from the national database being built at the Ministry of Health, to compare it with additional databases around the world using INFOODS and to turn it into a valid and accessible research tool. We hope that this work may aid other research groups in developing a computerised, nation-specific nutritional database and data-entry system adapted to their own specific local needs.

### Acknowledgements

We are indebted to Mr S. Daniel Abraham, founder of the International Center for Health and Nutrition, who supported this study. We also wish to thank the USDA, the Human Nutrition Center at the University of Texas and the European Prospective Investigation into Cancer and Nutrition for helping us develop our dietary tools. We are deeply grateful for the support we received in Israel: the Ministry of Health Department of Nutrition, the Central Statistics Bureau, Al-Outz University and Mrs Flora Lubin from the Gertner Institute.

### References

- Meir H, Reshef A, eds. *Food Composition Tables*. Jerusalem: Ministry of Health, Department of Nutrition, 1996.
- US Department of Agriculture (USDA). *Composition of Foods Raw, Processed and Prepared*. Agriculture Handbook No. 8-1-8-15. USDA Agriculture Research Service, 1976–87. Available at <http://www.nal.usda.gov/fnic/foodcomp/Bulletins/timeline.htm>.
- US Department of Agriculture (USDA). *Composition of Foods, SR12*. USDA Agriculture Research Service, Nutrient Data Laboratory, February 1999. Available at <http://www.nal.usda.gov/fnic/foodcomp/Data/SR12/sr12.html>.
- Guenther PM, Perloff BP, Vizioli TL. Separating fact from artifact in changes in nutrient intake over time. *J. Am. Diet. Assoc.* 1994; **94**: 270–5.
- Perloff BP. Recipe calculations for NFCS database. In: Murphy S, Rauchwarter D, eds. *Proceedings of the Tenth National Nutrient Data Bank Conference*. Springfield, VA: National Technical Information Service, Accession No. PB86-206281/HBF, 1985; 11–21.
- Powers PM, Hoover LW. Calculating the nutrient composition of recipes with computers. *J. Am. Diet. Assoc.* 1989; **89**(2): 224–32.
- Perloff B, Ahuja JKC. Development and maintenance of nutrient databases for national dietary surveys. *Public Health Rev.* 1988; **26**: 43–7.
- Food Intake Analysis System, Version 3*. University of Texas Health Science Center at Houston School of Public Health. US Department of Agriculture, Human Nutrition Information Service, 1996.
- Adams CF. *Nutrition Value of American Foods in Common Units*. USDA Handbook No. 456. Washington, DC: US Government Printing Office, 1975.
- US Department of Agriculture, Human Nutrition Information Service. *Manual of Food Codes for Individual Food Intake Survey, Release 2*. Springfield, VA: National Technical Information Service, Accession No. PB87-181020, 1986.
- Posati LP. Procedures for imputing values. In: Murphy S, Rauchwarter D, eds. *Proceedings of the Tenth National Nutrient Data Bank Conference*. Springfield, VA: National Technical Information Service, Accession No. PB86-206281/HBF, 1985; 124–33.
- McCance RA, Widowsom EM. *The Composition of Foods*, 5th ed. Royal Society of Chemistry and Ministry of Agriculture, Fisheries and Food. London: HMSO, 1992.
- Salvini S, Parpinel M, Gnagnarella P, Maisonneuve P, Turrini A. *Composition of Foods for Epidemiological Studies in Italy*. Milan: Istituto Europeo di Oncologia, 1998 (in Italian).
- Israel Ministry of Health, Department of Nutrition. *Weights and Portion Size Booklet*. Publication No. 217. Jerusalem: Ministry of Health, Department of Nutrition, June 2000.
- Block G, Dresser CM, Hartman AM, Carroll MD. Nutrient sources in the American diet: quantitative data from the NHANES II survey. II. Macronutrients and fats. *Am. J. Epidemiol.* 1985; **122**: 27–40.
- Block G, Norris JC. Source of energy and six nutrients in diets of low-income Hispanic-American women and their children: quantitative data from NHANES. *J. Am. Diet. Assoc.* 1995; **95**(2): 195–208.
- Hoover LW, ed. *Nutrition Data Bank Directory*, 7th ed. Columbia, MO: University of Missouri, 1988.
- US Department of Agriculture (USDA). *Food and Nutrient Intakes by Individuals in the United States by Sex and Age, 1994–96*. USDA Nationwide Food Surveys Report. USDA Agriculture Research Service, 1998.
- Favier JC, Ireland-Ripert J, Toque C, Feinberg M. *Repertoire Generale des Aliments. Table de Composition*, 2nd ed. Paris: Lavoisier TEC & DOC-INRA Editions, 1995.
- Health and Welfare Canada. *Nutrient Value of Some Common Foods*. Ottawa: Health and Welfare Canada, Health Services and Promotion Branch and Health Protection Branch, 1999.
- Verdier PC. The Canadian nutrient file. *J. Can. Diet. Assoc.* 1987; **34**: 523–8.
- Perloff B. Analysis of dietary data. *Am. J. Clin. Nutr.* 1989; **50**: 1128–32.
- Willett W, ed. *Nutritional Epidemiology*. New York/Oxford: Oxford University Press, 1998; 56.
- Perisse J. The heterogeneity of food composition tables. In: Hautvast JGAJ, Klaver W, eds. *The Diet Factor in Epidemiological Research*. Euronut Report 1. Wageningen: 1982; 100–5.
- Rand WM. Food composition data: problems and plans. *J. Am. Diet. Assoc.* 1985; **85**: 1081–3.
- Schlotke F, Becker W, Ireland J, Møller A, Ovaskainen M-L, Monspart J, et al. EUROFOODS Recommendations for Food Composition Database Management and Data Interchange. COST Action 99 – EUROFOODS, 2000.
- LanguaL 2000: Introduction and Users' Manual*. COST Report EUR19542. Luxembourg: European Commission, 2000.
- Miller A, Ireland J. *LanguaL 2000: Thesaurus*. COST Report EUR19540. Luxembourg: European Commission, 2000.