DEVELOPING AN INTERNATIONAL FOOD COMPOSITION DATABASE: AN ILLUSTRATION FROM THE NUTRITION CRSP

J.A. Ricci,1 N.W. Jerome,2 W.A. Moussa,3 O. Galal,4 G.G. Harrison,5 A. Kirksey6

INTRODUCTION

Many types of research on human dietary intake rely on food composition tables to convert quantitative data on food and beverage consumption to the caloric and nutrient content of a diet. For this reason, the appropriate food composition database must be identified and applied to the research.

Scientists conducting research on United States populations have access to a large network of food composition databases. For them, the selection issue often depends only on the appropriateness of the database for testing particular scientific hypotheses. The complexity of this process can increase significantly in international research. The required approach can range from the minor modification of a preexisting database to the complete construction of local, regional, or national food composition tables. This paper highlights the approach used to develop a food composition database for Egypt.

In 1981, the Office of Nutrition, United States Agency for International Development (USAID) provided funds to support the Nutrition Collaborative Research Support Program (CRSP).7 The research program investigated cross-culturally the effect of chronic mild-to-moderate undernutrition on human functional outcomes believed to shape the development of individuals, households, and communities.

Since the Nutrition CRSP aimed to generate information that could be generalized to most of the less developed world, the program implemented a "parallel projects" approach in three countries with different biological and sociocultural ecosystems. The three selected field sites were Kalla, Egypt; the Salis Valley, Mexico; and Embu District in Eastern Province, Kenya.

Each of the three country projects posed an identical core of key research questions. However, methodologic procedures to answer these questions differed among projects to accommodate environmental characteristics, resources, and sociocultural contexts unique to each field situation. For example, the country-specific food composition database developed for Egypt had to take into account the specific foods and beverages of Egypt, Egyptian food preparation and preservation methods, current food composition tables and earlier attempts at developing a local food composition database.

1 The Johns Hopkins University School of Hygiene and Public Health, Baltimore, MD
2 United States Agency for International Development, Washington, D.C. and University of Kansas School of Medicine, Kansas City, KS
3 Institute of Nutrition, Cairo, Egypt
4 National Research Center, Cairo, Egypt
5 University of Arizona, Tucson, AZ
6 Purdue University, West Lafayette, IN
7 The Collaborative Research Support Program on Food Intake and Human Function was carried out under USAID grant no. DAN-1309-G-SS-1070-00.
DESCRIPTION OF THE EGYPT PROJECT STUDY SITE

Kalama, Egypt, is a periurban community of approximately 10,000 individuals situated in the Nile River delta just 25 kilometers north of downtown Cairo. Kalama’s climate is hot and humid, and rainfall is extremely sparse falling no more than six times per year. There are three distinct seasons - the very hot, humid and dusty summer from April to August, the mild autumn from late September to November and the cool winters from December to March.

In 1982, over 50 percent of Kalama’s predominantly Moslem population was under 21 years of age. An average household consisted of approximately six individuals. Two-thirds of all village households were nuclear families, consisting of a married couple and their unmarried offspring; the remaining one-third were extended family households. Most older village women have little or no formal education. However, recent trends indicate that norms of female education are changing as the younger generation of females are strongly encouraged to attend school.

Agriculture is the basis of economic life. However, employment in the non-agricultural sector is continuously increasing particularly for the younger generations.

THE EGYPT NUTRITION CRSP FOOD COMPOSITION DATABASE

The Egypt Nutrition CRSP Food Composition Database contains information on energy and 15 nutrients for over 1,000 unique entries. The food composition database, assembled by food groups, includes the following information for each food or beverage item:

1. Name and brief description;
2. A four-digit food identification code;
3. All applicable two-digit “state of food” codes;
4. An edible portion factor for converting food intake data recorded in grams “as purchased” into grams of “edible portion” consumed;
5. The energy content in the edible portion of 100 grams of each dietary item “as acquired” or “as purchased”;
6. Fifteen food components in 100 grams of the edible portion of each dietary item; refer to Table 1;
7. Reference number indicating the source of the food composition data.

Information presented subsequently in this paper reviews the method for developing the Egypt Nutrition CRSP Food Composition Database. The review proceeds from general to specific and concludes with a discussion of the potential errors associated with its use.

GENERAL APPROACH TO DEVELOPING AN INTERNATIONAL FOOD COMPOSITION TABLE

Studies on human dietary intake require culturally appropriate and sensitive research methods. This principle also applies to the construction of locally relevant international food composition tables.

The general approach requires five major steps. The first step involves identifying and evaluating available unpublished in-country and published food composition tables applicable to the country.

The second major step entails establishing familiarity with the environmental and sociocultural context of food consumption. This is accomplished by conducting focused ethnographic and social survey research to gather information on 1) the typical dietary pattern identified by both the local and scientific names of foods, 2) factors which alter the typical dietary pattern, 3) seasonal availability of plant and animal foods, 4) household food acquisition strategies, 5) community-specific measures of quantifying foods and beverages, 6) standard
measures of quantifying foods and beverages consumed, and 7) local cuisine and culinary practices including traditional food processing and preparation.

Step three utilizes the ethnographic and social survey data to construct a food coding system based on the local dietary pattern. The system must incorporate all known locally-consumed foods and beverages and accommodate expansion as changes in food habits occur. Each food code or identification number must be unique to accurately link food consumption data with corresponding food composition data during the computerized conversion of foods and beverages to energy and nutrients.

The fourth step involves organizing the database. Here, food composition data gathered from the field, select laboratory analyses such as those for composite diets, and published tables of food composition come together in a format which reflects both the dietary practices of the community and the specific needs of the scientific investigation.

The fifth and final step checks the completed database for internal integrity and validity. Arithmetic or data entry errors can be detected and corrected, while other less obvious sources of error can be identified for discussion when reporting study results.

CREATING A FOOD COMPOSITION DATABASE FOR EGYPT

Experiences from the Egypt Nutrition CRSP fellow. We illustrate how the five-step approach was applied to develop an international food composition database.

STEP:1 Reviewing and Evaluating Available Food Composition Tables

United States and Egyptian scientists reviewed and evaluated Egyptian and other Middle Eastern food composition tables and collaboratively decided to create a food composition database specifically for the Egypt Nutrition CRSP. Working within financial and technologic constraints, four major steps were taken to develop the Egypt project’s food composition database.

STEP:2 Obtaining Qualitative and Quantitative Information on the Sociocultural Context of Food Consumption in Kalama

The Kalama Dietary Pattern

A food frequency protocol in combination with qualitative observation on local food consumption provided data on the Kalama dietary pattern. The dietary items obtained from this methodologic approach formed the core of the project’s food composition database.

The Kalama dietary pattern mirrors the traditional rural Egyptian diet. Bread, the dietary staple, is consumed daily by virtually all individuals and accounts for most of their daily caloric intake. Other principal components of the dietary pattern include; sugar, fava beans (or broad beans), home-made cheese from goat’s or buffalo’s milk, rice, and stewed vegetables available seasonally. Meat and poultry are consumed infrequently and fish, rarely.

Factors Altering the Dietary Pattern

Significant alterations to the typical dietary pattern occur during frequent Moslem religious observances. During these periods the typical dietary pattern expands to include a variety of festive mixed dishes, sweets, meat, ceremonial fruits, and nuts.

Household recipes for specially prepared foods were obtained during holidays (and also periods of usual consumption). This procedure increased the accuracy in estimating dietary intake by accounting for
interhousehold intracommunity variability in food preparation methods. The Egypt project identified 11 dietary items for which standardized recipes were obtained.

Seasonal Availability of Dietary Items

Knowledge of the local agricultural cycle helps to predict the seasonal market availability of a variety of locally produced fruits and vegetables. Information on seasonal diversity of foods ensures that the food composition database captures the entire range of items in the annual diet.

In Kalama, the primary agricultural crops include maize, beans, parsnips, carrots, leaf crops, and some other vegetables. In the winter citrus fruits are harvested. Fresh dates are available in the late summer and autumn and in the summer, tropical fruits such as watermelon, guava, and mangoes are widely available.

Household Food Acquisition Strategies

Knowledge of food acquisition strategies helps in designing a food composition database since differences in commercial preparation and processing techniques introduce the potential for variation in food composition. In Kalama, households acquire dietary items from both commercial and home-based sources.

When commercially prepared foods play a significant role in the dietary pattern, their food composition should be determined by laboratory analyses to avoid potential errors resulting from imputed data. This will improve the database's accuracy in estimating individual energy and nutrient intakes.

Standard Measures of Quantifying Foods and Beverages

The Egypt Nutrition CRSP relied on ethnographic data to develop culturally-appropriate food intake research methods. Similarly, the format of the food composition database must be compatible with food intake data. Two basic format decisions were made.

The first concerns the use of local household measures versus metric weights to record dietary items. The Egypt Nutrition CRSP decided that food intake data should be obtained in household measures and then converted to metric. This decision was based on the variability of food preparation and measuring techniques in village households and the local professional tradition in obtaining dietary data. Quality control measures were applied rigorously to ensure accuracy in the measurement and conversion process. Therefore, the food composition database reports the food composition of 100 grams of each dietary item.

The second decision concerns the use of "as purchased" versus "edible portions" of dietary items. In Kalama, the traditional style of communal eating from a common serving dish makes it difficult, if not impossible, to precisely quantify individual food intake. Therefore, data collectors recorded quantities of food "as acquired" for preparing a given dish for family consumption, i.e. quantities of ingredients were recorded in the "as purchased" state. On the other hand, commercially-prepared foods were recorded in the "edible portion" state. To accommodate the data collection protocol, the project added a "edible portion" factor to the food composition database to convert grams of foods "as purchased" to grams of "edible portion."

Aspects of Cuisine: Traditional Food Preparation and Processing

Ethnographic research identified a range of food preparation and processing methods in Kalama permitting women without access to the convenience of refrigeration to safely store, preserve, and prepare foods in a hot dry climate. These were taken into account in estimating nutrient content since food preparation and processing can alter food composition.

In Kalama, many foods such as fruits, vegetables, and cheese are preserved by pickling or salting. In addition, women traditionally cook foods for long periods of time; "stewed" vegetables are very common. Cooked
dishes prepared in two-day quantities are periodically reheated to prevent spoilage and the "next day portion" of day-old prepared dishes are reheated up to the boiling point just before eating.

Due to these realities, the Egypt Nutrition CRSP developed a "state of food" coding system to accompany the basic food identification code to specify food preparation and preservation methods. Step 3 provides more information on the food coding system.

**STEP 3: Developing the Food Coding System**

The need to specifically determine the energy and nutrient content of a food item led to the development of a sophisticated yet simple four-digit food coding system. This system linked food consumption data with the corresponding food composition data during the computerized conversion of foods to energy and nutrients.

The four-digit food code has three main features: 1) the ability to accommodate existing foods and beverages while permitting future expansion within the system, 2) a system of grouping foods based on traditional food groups and frequency of consumption within food groups, and 3) a complementary two-digit "state of food" coding system to specify both preparation method and "technologic" state, i.e. the "state" of food as consumed. Refer to Figure 1.

The first two digits of the four-digit food code represent the food group, e.g. cereals/cereal products. The second two digits of the four-digit food code denote the specific member of the food group, e.g. rice.

The two-digit state of food codes even further describe the food or beverage item to facilitate a more accurate estimation of its energy and nutrient content. Examples of states of food are boiled, stewed, and raw.

**STEP 4: Organizing the Egypt Nutrition CRSP Food Composition Database**

The core of the Egypt Nutrition CRSP Food Composition Database was obtained from three published tables of food composition. Additional information was obtained from the Egypt Nutrition Institute's unpublished food composition tables, nutrient labels supplied by manufacturers of commercially prepared products, proximate analyses performed by two United States laboratories, and other imputed data derived from standardized local recipes. Integrating these data from several sources while ensuring accuracy and consistency in the final composite product demanded a systematic approach and an eye for detail.

**Data from Published Tables of Food Composition**

Information on food energy and other nutrients were carefully and systematically selected from three primary reference documents: *Food Composition Tables for the Near East (FAO)*, *Food Composition Tables for Use in East Asia (FAO)*, and *Bowes and Church's Food Values of Portions Commonly Used.*

When nutrient information on cooked foods was unavailable, vitamin and mineral losses from cooking were estimated and nutrient values were reduced accordingly. The publication, *Provisional Table on Percent Retention of Nutrients in Food Preparation (USDA)*, provided standardized percents reflecting proportion of nutrients retained in different foods after cooking by a variety of methods. Knowledge of local culinary practices were invaluable in applying data from this document.

Data from Laboratory Analysis

Two United States laboratories provided data on the energy, available carbohydrate, fat, and protein content of commonly consumed "fast foods" in Kalama. To maintain consistency with the rest of the database, total carbohydrate was estimated. Thus, laboratory results on available carbohydrate were then converted to total carbohydrate by adding grams of dietary fiber to grams of available carbohydrate.
Data from Recipes

The energy and nutrient composition of eleven Egyptian composite or mixed dishes were derived from local recipes in a systematic way. Information on the energy, protein, fat, and total carbohydrate content of each ingredient in a recipe was obtained from the Egypt Nutrition CRSP Food Composition Database. The sum of the energy and nutrients of all ingredients equalled the recipe's total.

Next, the total weight of the ingredients in a recipe was adjusted to reflect net changes in weight and nutrient content per unit weight after cooking or baking. Examples include the increase in caloric density associated with a decrease in weight from moisture loss during baking or an increase in weight from fat absorption during frying. The publication, *Food Yields Summarized by Different Stages of Preparation* (USDA3) provided data on proportional weight changes.

Finally, the energy, protein, fat and carbohydrate content of 100 grams of the prepared recipe was calculated by a simple proportionate adjustment.

**STEP 5: Checking the Food Composition Database for Internal Integrity and Validity**

After the Egypt project’s food composition database was assembled, manually checked, and double entered, an algorithm checked the database for internal inconsistencies. The Atwater energy equivalent factors were applied to the values for protein, fat, and total carbohydrate to calculate approximate energy content for each dietary item. The calculated energy value was compared to the database energy value and differences greater than five percent were flagged, checked, and corrected as necessary.

The Egypt Nutrition CRSP food composition database was validated against the results of laboratory analysis. Seven Kalamans of different age, sex and physiologic state (females only) provided individual food consumption data and concomitant 24-hour composite dietary samples for the validation protocol. For each of the seven individuals, the energy, protein, fat and total carbohydrate values derived from the individual food intake data using the food composition database were compared to the results of proximate laboratory analysis. The percentage difference is expressed relative to laboratory values. The results indicate a mean difference ranging from 7.9 percent for carbohydrate to 70.3 percent for fat. Refer to Table 2.

**SOURCES OF ERROR IN ESTIMATING DIETARY INTAKE**

Despite the care in developing food composition databases, users of food composition tables should be aware of potential errors and inconsistencies in the data. There are many reasons for this including the natural nutrient variability, diversity in food preparation techniques even when recipes are used, and errors derived from the laboratory methods used to determine the energy and nutrient content. Additional sources of error are often introduced in developing countries due to inherent technologic, cultural, and economic constraints in conducting certain types of research in this setting. Errors in the Egypt Nutrition CRSP data base stem from such factors as an inability to completely account for the variability in traditional food preparation and the inclusion of both home-prepared and commercially-prepared mixed dishes in the dietary pattern. These problems can be resolved with advances in research.

Finally, food composition data permit an estimate of only the total level of energy and nutrients in dietary items. Critical factors in determining dietary adequacy such as actual digestibility and bioavailability of nutrients certainly extend beyond the capacity of any food composition table. Even though the specific Atwater energy equivalent factors adjust for digestibility, diets which are relatively high in dietary fiber (as are many in the tropics and semi-tropics) may not be as well-digested as the factors predict (Murphy4).

The reduced bioavailability of some nutrients poses a special problem in these settings. For example, the major source of dietary iron for some populations may be from plant sources with little bioavailable iron. In addition, complementarity with other dietary items especially those high in some types of fiber and phytates
may reduce both the digestibility of the diet and the bioavailability of some nutrients (e.g. zinc). To further compound the problem, the high prevalence of diarrheal disease can also significantly reduce the absorptive capacity of the intestinal mucosal surface in individuals already at risk to malnutrition.

CONCLUSIONS

This paper reviewed the conceptual and procedural approaches for developing an international food composition database using a recent Egyptian case as an example. The limitations of food composition tables have also been discussed here and by others (Watt;7 Rand;8 Wu Leung;9 Hertzler and Hoover;10 Mayer11). Please note however, that food composition tables are not the only potential source of error in the estimation of dietary intake. The validity and reliability of methods to measure dietary intake are well-documented and also warrant consideration when planning studies and interpreting dietary information (Brown;12 Block;13 Walker and Blettner14).

Finally, the development of food composition tables is never static. It requires continuous expansion, reorganization, update, and revalidation; however, despite these precautions nutrient information on a particular dietary item may be either incomplete or entirely lacking. Our responsibility in these circumstances is to make rational well-documented decisions when filling in the blanks until the time when adequate sampling and standardized laboratory analysis can provide more representative, valid and reliable substitutes for this imputed information.

REFERENCES


6. Murphy, S. Personal communication, 1989.


---

**Figure 1.** Food Coding System. *Egypt Nutrition CRSP.*
### Table 1
Food components in the Egypt Nutrition CRSP Food Composition Database.

<table>
<thead>
<tr>
<th>FOOD COMPONENT</th>
<th>UNIT(S) OF PRESENTATION</th>
</tr>
</thead>
<tbody>
<tr>
<td>Energy</td>
<td>kilocalories</td>
</tr>
<tr>
<td>Water</td>
<td>grams</td>
</tr>
<tr>
<td>Protein (animal sources)</td>
<td>grams</td>
</tr>
<tr>
<td>Protein (plant sources)</td>
<td>grams</td>
</tr>
<tr>
<td>Fat (animal sources)</td>
<td>grams</td>
</tr>
<tr>
<td>Fat (plant sources)</td>
<td>grams</td>
</tr>
<tr>
<td>Total carbohydrate</td>
<td>grams</td>
</tr>
<tr>
<td>Calcium</td>
<td>milligrams</td>
</tr>
<tr>
<td>Iron (animal sources)</td>
<td>milligrams</td>
</tr>
<tr>
<td>Iron (plant sources)</td>
<td>milligrams</td>
</tr>
<tr>
<td>Sodium</td>
<td>milligrams</td>
</tr>
<tr>
<td>Thiamin</td>
<td>milligrams</td>
</tr>
<tr>
<td>Riboflavin</td>
<td>milligrams</td>
</tr>
<tr>
<td>Niacin</td>
<td>milligrams</td>
</tr>
<tr>
<td>Vitamin A (retinol)</td>
<td>milligrams</td>
</tr>
<tr>
<td>Vitamin A precursor (B-carotene)</td>
<td>micrograms, retinol</td>
</tr>
<tr>
<td>Ascorbic acid</td>
<td>milligrams</td>
</tr>
<tr>
<td>Crude fiber</td>
<td>grams</td>
</tr>
<tr>
<td>Ash</td>
<td>grams</td>
</tr>
</tbody>
</table>

### Table 2
Summary of percentage differences between food composition database and laboratory analysis.

<table>
<thead>
<tr>
<th>Sample</th>
<th>Percent differences</th>
</tr>
</thead>
<tbody>
<tr>
<td>Adult male</td>
<td>Protein  Fat  Carbohydrate Kcal⁺</td>
</tr>
<tr>
<td>Non-pregnant non-lactating</td>
<td>10.2  97.7  -5.9  13.9</td>
</tr>
<tr>
<td>Pregnant 9 months</td>
<td>18.7  57.0  2.3  19.1</td>
</tr>
<tr>
<td>Lactating 7 months</td>
<td>-5.6  11.5  1.3  2.2</td>
</tr>
<tr>
<td>Pregnant 4 months and</td>
<td>63.9  252.7  59.6  91.4</td>
</tr>
<tr>
<td>lactating 15 months*</td>
<td></td>
</tr>
<tr>
<td>School-age child</td>
<td>24.1  63.9  28.0  31.3</td>
</tr>
<tr>
<td>Toddler</td>
<td>-7.2  -2.5  -16.1 -12.1</td>
</tr>
<tr>
<td>Mean difference</td>
<td>13.2  70.3  7.9  19.5</td>
</tr>
</tbody>
</table>

⁺Difference is expressed relative to laboratory values.

*Kcals determined using Atwater energy equivalent factors.

*Possible error in sample preparation for analysis.