# 16th National Nutrient Databank Conference

*NUTRIENT DATABASES FOR THE 1990·s:  EXCELLENCE IN DIVERSITY*

June 17-19, 1991
San Francisco, California

- Effects of Nutrient Data Changes on Results on NFCS 1977-78 and NFCS 1987-88
  Patricia M. Guenther and Betty P. Perloff

Updates

- Update of HNIS Survey Activities
  Alanna J. Moshfegh
- USDA Nutrient Data Base Update
  Betty P. Perloff
- NHANES III Update
  Robert J Kuczmarski and Margaret McDowell
- Status of INFOODS Publications
  William Rand and John Klensin
- Guide for Nutrient Data Users
  Grace J. Petot

Time-Related Databases

- Maintaining Time-Related Databases for Dietary Data Collection and Nutrient Calculation
  I. Marilyn Buzzard

Obtaining and Using Industry Data

  - Sandra J. Bartholmey
  - Jenny Chin
  - Frances H. Seligson
  - Mary Stevens

Services for Databank Users/Developers

- Databases for Shelf Labeling: "GTI Data Docuement Systems"
  KarenFalk
- Using the Menu Census Survey to Estimate Dietary Intake: Post Market Surveillance of Aspartame
  I. Jack Abrams
- Langual for Database Users
  Elizabeth C. Smith
- Using Electronic Mail (EMAil..) for Database Communications
  Lorry Scura and John C. Klensin
- Addressing and Related Issues in Electronic Mail Systems
  John C. Klensin and Lorry Scura

Concurrent Session II

- Food Frequency Data Entry and Analysis Program
  R. Sue McPherson
- Nutrient Databases for Food Frequency Instruments
  Laura Sampson
- Current HNIS Nutrient Data Research
  Ruth Matthews
- Management of Food Composition Databases in Foodservices Settings
  Loretta W. Hoover
- Maintenance of Food Composition Databases
  Sally Schakel and Brian Westrich
- Management of Food Composition Data Bases
  Brucy C. Gray

Using International Nutrient Data:

# ACKNOWLEDGEMENT

The Nutrient Databank Conference is made possible in part by generous financial support of the following institutions and companies:

Human Nutrition Information Service, USDA

Best Foods, CPC International Inc.

California Beef Council

Campbell Soup Company

Carnation

The CBORD Group, Inc.

Church and Dwight

The Coca-Cola Company

Frito-Lay, Inc

Graphic Technology, Inc.

Hershey Foods Corporation

Hoffmann - LaRoche, Inc.

International Network of Food Data Systems

Kraft General Foods, Inc.

Nabisco Biscuit Company

The Procter and Gamble Company

# Foreword

The National Nutrient Databank Conference has been held annually for the past 16 years. It is organized by a group of interested volunteers who give generously of their time. A steering committee and four working committees carry on the functions of organizing the conference and communicating conference activities. Interested persons are invited to volunteer for committee assignment at each conference. The continued success of these conferences is due in large part to the work and dedication of these volunteers. Committee chairs for the 1991 conference were: Steering Committee, **Al Riley** of Campbell Soup Company; Program Committee, **Jean Pennington** of the Food and Drug Administration; Communications Committee, **Betty Perloff** of USDA-HNIS; Database Committee, **Jack Smith** of the University of Delaware; Local Arrangements Committee, **Suzanne Murphy** of the University of California, Berkeley.

Proceedings have been published following most of the conferences, and many are still available for purchase. A listing of proceedings publishers is given in the back of this issue. Special thanks go to **The CBORD Group, Inc.,** who have donated the cost of printing and distributing the proceedings to all conference participants for the past six conferences. We also appreciate the assistance of **Cici Hyde** at the University of California, Berkeley, and **Lisa Morse** at the CBORD Group, Inc., for their help in preparing the manuscripts for this publication.

The conference organizers wish to thank the **U.S. Department of Agriculture** and this year's corporate and institutional sponsors, listed earlier in these proceedings, for their generous support of the conference. Special recognition goes to **USDA's Human Nutrition Information Service,** not only for their financial support but also for the many contributions made by their professional staff.

The 1992 conference will be held on June 8-10 in Baltimore. Information about that conference is available by contacting **Jack Smith,** Department of Nutrition and Dietetics, Alison Hall, University of Delaware, Newark DL 19716.

# Committee List

## NATIONAL NUTRIENT DATABANK COMMITTEES

### STEERING COMMITTEE

Al Riley, Campbell Soup Company, Chair

Suzanne Murphy, University of California, Berkeley
Jean Pennington, Food and Drug Administration
Betty Perloff, Human Nutrition Information Service, USDA
Jack Smith, University of Delaware
Bob Rizek, Human Nutrition Information Service, USDA

### LOCAL ARRANGEMENTS COMMITTEE

Suzanne Murphy, University of California, Berkeley, Chair

Rita Mitchell, University of California, Berkeley
Mary Anne Burkman, Dairy Council of California
Ann Coulston, Stanford University Hospital
Pat Crawford, University of California, Berkeley
Doris Dare, San Francisco General Hospital
Joanne Ikeda, University of California, Berkeley
Kathleen Kelley, University of California, Berkeley
Mary J. Kretsch, Western Human Nutrition Research Center
Kathryn Sucher, San Jose State University
Karen Todd, University of California, San Francisco

### COMMUNICATIONS COMMITTEE

Betty Perloff, Human Nutrition Information Service, USDA, Chair

Maggie Clausell, University of Maryland
Judi Douglass, TAS, Inc.
Lois Fulton, Human Nutrition Information Service, USDA
Loretta Hoover, University of Missouri
Joanne LaValle, The CBORD Group
Won Song, Michigan State University

### DATABASE COMMITTEE

Jack Smith, University of Delaware, Chair

Carol Ballew, Community Health Research
Beverly Benes, Northern Illinois University
Marilyn Buzzard, University of Minnesota
Charlene Hamilton, Framington State College
David Haytowitz, Human Nutrition Information Service, USDA
Linda Hicks, University of Texas
Catherine C. Johnson, Pennington Biomedical Research Center
Beth Lucey, The CBORD Group
Grace Petot, Case Western Reserve University
Phyllis Stumbo, University of Iowa

### PROGRAM COMMITTEE

Jean Pennington, Food and Drug Administration, Chair

Marilyn Buzzard, University of Minnesota
Patty Deuster
Connie Dresser, National Cancer Institute
Joanne Holden, USDA-ARS-BHNRC
John Klensin, Massachusetts Institute of Technology
Molly Kretsch, USDA-WHNRC
Sue McPherson, University of Texas
Suzanne Murphy, University of California at Berkeley
Linda Posati, USDA-FIAD
Lorraine Scura, USDA-HNRC
Jack Smith, University of Delaware
Won Song, Michigan State University
Phyllis Stumbo, University of Iowa

# Introductory Remarks

**Doris Howes Calloway**
**University of California, Berkeley**
**Berkeley, California**

It is an unexpected pleasure to find myself welcoming you on behalf of our department chair, Janet King, and the University of California to the 16th National Nutrient Databank Conference. Dr. King regrets that she is unable to be here and sends her best wishes for a rewarding and pleasurable conference.

Research on the composition and nutrient content of foods has been featured in the land-grant universities from their very beginning. Since passage of the Hatch Act in 1887 much of this research has been in partnership with the United States Department of Agriculture. Our joint sponsorship of this conference is most appropriate.

It might seem that we ought to know all there is to know about the composition of U.S. foods by now. But food supplies and consumption patterns change. Also, it seems the more we know, the more there is to know.

I once thought I must know all there was to know about collard greens. Among the many jobs that paid my way through Ohio State was a laboratory assistanceship in the Experiment Station. Why collard greens was the focus of research escapes me, but it was a large project with many variables: composition of plants harvested at different stages, from opposite ends of rows in plots, stored and prepared in different ways, etc. I learned that all of these variables affected nutrient content--a useful perspective to have acquired early in our profession.

A decade or so later, Brassica vegetables were on my plate again. Our research group at the Quartermaster Food and Container Institute for the Armed Forces followed up early French work showing that guinea pigs fed a diet of oats and bran with cabbage were more resistant to radiation injury than were animals given the same diet with beets. We confirmed the French study and found that broccoli was more effective than cabbage. Unsnarling the known nutrient variables--vitamins A and C were important, for example--was a major task. Our last attempt to mimic the effect of broccoli with known compounds involved 30+ pure substances; the mix neither looked nor smelled like broccoli and it didn't affect radiation injury either. We now know much more about the properties of this family of plants, including the fact that their consumption is associated with reduced risk of colon cancer. We still don't know exactly why. The point or points of this story are: that foods and nutrients are not the same thing; that foods grouped withing a class have some but not all properties in common, or not to the same extent; that discovery of previously unrecognized physiological properties results in a need to know more about old substances--e.g., vitamin C and beta-carotene--and especially, a need to identify and quantify other factors--almost none of which will prove to be nutrients as classically defined but are delivered willy-nilly along with nutrients.

There is nothing I could add to this groups knowledge about the impact of the changing U.S. food supply and intake patterns on data base needs. It used to be an article of faith that food habits are nearly immutable--hard to believe even then, I would have thought, given the example of the worldwide response to Coca Cola. A small canvass of San Francisco restaurants--which I hope you will manage to fit into your busy conference schedule--will illustrate, in case there is any need for it, that the U.S. is a rich, diverse and increasingly globalized village.

This conference is testimony to your commitement to expansion and improvement of the national, and now international, composition information base. We wish you well in this important task.

# Nutrient Databases for the '90's

**Elaine R. Monsen, Ph.D., R.D.**
**University of Washington, Seattle**
**Seattle, Washington**



I appreciate the opportunity of meeting with you to discuss the important issues of nutrient databases. I will look at the Topic of Excellence in diversity in the following four sections: diversity of food components; diversity of the use of databases; diversity of issues to be considered; and finally, diversity of application and research needs.

**Diversity of Food Components**

Food contains tens of thousands of components. The various components of food (Table 1) include those that naturally occur within foods, those that are added to foods intentionally or unintentionally, and engineered and designer foods. The databases generally emphasize the energy-providing nutrient and essential nutrients with some data for dietary fiber. In general, only 23-90 of the thousands of food components are provided by nutrient databases. The other naturally-occurring components, natural toxins and anti-nutrient substances, and other naturally-occurring compounds such as enzymes, sterols, pytochemicals, and organic acids, are infrequently, if ever displayed. Caffeine is an exception and is found in several of the databases.

The largest number of components of foods are the additives and contaminants that are either intentionally and inadvertently incorporated into foods. A very small list of examples are included in Table 1. There is going to be continuing interest in the components of foods that may impact physiological responses to food consumption. Public interest has already been expressed about agricultural contaminants and inadvertent additives such as leached metals, iodide, and migrated packaging materials. Other compounds that are formed in food processing, storage and preparation also are of interest. Some of these are flavorants. Many compounds, however, require further investigation to

**Table 1 - COMPONENTS OF FOODS**

**I. Naturally occurring components**

A. Energy providing nutrients
   e.g. Carbohydrate, Protein, Fat, Alcohol

B. Essential nutrients
   e.g. Vitamins, minerals, amino acids, fatty acids, water

C. Dietary fiber
   e.g. pectin, gums, hemicellulose, cellulose, ignins

D. Natural toxins and anti-nutrient substances
   e.g. goitrogens, avidin

E. Other naturally occurring compounds
   e.g. DNA, RNA, enzymes, sterols, caffeine, nonessential minerals, nonessential amino acids, bioflavanoids, indoles, pytochemicals, flavorants, organic acids

**II. Additives and contaminants**

A. Intentional additives
   e.g. NaCl, preservatives, stabilizers, flavorants, artificial sweeteners, nutrient supplements

B. Agricultural contaminants
   e.g. pesticides, herbicides, growth hormones, antibiotics

C. Microbial and fungal toxins, and parasites
   e.g. aflatoxin and other mycotoxins, salmonella enterotoxin, and giardia lamblia

D. Inadvertent additives
   e.g. leached metals, iodide, migrated packaging materials

E. Compounds formed in food processing, storage, and preparation
   e.g. congeners in wines and spirits, compounds formed in charcoal broiling such as formaldehydes, furfurals, benzenes, free radicals, singlet oxygen fermentation products such as lactic acid, trans fatty acids from hydrogenation of oils

**III. Engineered and designer foods**
   e.g. fat substitutes, egg substitutes, protein isolates and concentrates

ascertain their ultimate compatibility in human physiology; particularly the free radicals and singlet oxygen, trans fatty acids produced from hydrogenation of oils, and compounds that are formed when proteins are heated to a high temperature e.g. formaldehydes, fufurals, and benzenes.

The last category is engineered and designer foods of which we are undoubtedly going to see more. At this point, most of these foods are designed as substitutes for naturally occurring foods or food components such as fat or eggs, or vegetable protein isolates formulated to replicate animal protein products. With further refinement of dietary needs we will find additional engineered and designer foods produced.

**Table 2 - Diversity of use of nutrient databases in the 1976-1991 April issues of**

**THE JOURNAL OF THE AMERICAN DIETETIC ASSOCIATION**

| | |
|---|---|
| 1976 | Children with Prader-Willi Syndrome |
| | Fatty acids in cereal foods |
| 1977 | Consumers of fast food meals |
| | Nutrient intake in hyperactive children |
| 1980 | Nutrient intake of preschool vegetarian children (includes 11 amino acids) |
| 1981 | Elderly participants in congregate meal programs |
| | Nutrition education for nursing home residents |
| 1982 | Dietary fiber consumption in a Dutch population |
| 1983 | Alcohol intake in NFCS |
| | Food purchased for nursing homes |
| | Lacto-ovo vegetarians |
| 1984 | Dietary screening device for clinics |
| 1985 | Dietary intervention in hypertensives |
| | Patients with diabetes receiving insulin pump therapy |
| | Nutrient content of published weight reducing diets |
| 1986 | Family-based behavioral weight control programs for obese children |
| | Hospitalized patients with eating disorders |
| | Beverages in the diets of American teenagers |
| | Dietary characteristics of hyperactive boys |
| 1987 | Urban homebound population |
| | Patients with Crohn's disease |
| | Commercial eating and nutrient adequacy |
| | Urea kinetic modeling and nutritional management of patients undergoing hemodialysis |
| 1988 | Trace element status of children with PKU |
| | Caffeine intake of children in Bogalusa heart study |
| | Chloride distribution in foods |
| 1990 | Bulimic female college students |
| | Culturally diverse low income pregnant women |
| | Adult obese women: CPF intake compared with blood values and weight loss |
| | Food consumption and dietary goals of healthy volunteers |
| 1991 | Diet and serum lipids in Vegans |
| | Disabled residents in long-term care facility |

## Diversity of Use of Nutrient Databases

To gain some insight into the wide use that databases have been employed, I looked through the April issues of the *Journal of the American Dietetic Association* from 1976, the year of the first nutrient database conference, to 1991. In these issues, 32 articles utilizing nutrient data were published. The wide variety of application is remarkable (Table 2), ranging from nutrient intake, food and alcohol intake, fiber intake, of healthy individual; of subjects with specific diseases such as Prader-Willi Syndrome or Crohn's disease; to populations with specific food philosophies such as vegans and vegetarians. The uses of nutrient databases cross age and gender lines.

Dividing the 16 years of April issues into four groups of four years I tallied the nutrients that were presented (Table 3). It is fascinating to note that from 1976-79 only nine nutrient components were presented in the four articles published in April of that four year span. By 1980-83 one sees the impact of the first release of software programs for nutrient databases as the number of nutrients jumps to 21, plus 11 amino acids in one report. The seven articles during that period included several more vitamins and minerals, as well as, dietary fiber, pectin and alcohol. In the next four year span, 1984-87, 12 articles were published with a total of 24 nutrients presented. One saw addition of minerals and of sugar and the beginning of the interest in fatty acids with the inclusion of saturated fatty acids (SFA) and polyunsaturated fatty acids (PUFA). The last four year period, 1988-1991, tallied nine published articles in April issues and the addition of mono-unsaturated fatty acids (MUFA) along with SFA and PUFA and the addition of caffeine.

Several observations are apparent from scanning April issues from 1976-1991. First, there is greater diversity in food selection than the databases accommodate. Secondly, there is greater variability in food composition than databases accommodate. Third, only a limited number of nutrients are estimated. Fourth, other components of foods are neither estimated nor considered. Fifth, there is infrequent information regarding missing or imputed values. One striking observation is that as soon as nutrient database programs were available to researchers, they were quickly utilized, as is readily apparent from the jump from 9 to 32 nutrients reported between 1976-79 to 1980-83 with the release of nutrient data programs.

## Diversity of Issues

*Queries to software providers -*

To obtain information regarding the diversity of programs available to consumers and researchers, I checked with several of the major providers and asked several questions. The answers indicated that nutrient data software programs became available in the early 1980's: 1980, 1982, 1984, etc. The number of foods included in the nutrient databases ranged from 2,000 to 16,000.

The number of nutrients or components ranged from 25 to 94. The nutrient database with the fewest components included 3 related to protein (total protein, animal protein, plant protein), 2 related to carbohydrates (complex carbohydrate and added sugar), 8 related to fat (total fat, SFA, MUFA, PUFA, animal fat, fish fat, plant fat, and cholesterol). Five minerals were included (sodium, potassium, iron, calcium, and phosphate), 2 vitamins (vitamin A and C), 3 other components (alcohol, caffeine, and aspartame). With the addition of energy and fiber a total of 25 nutrient components were presented. Another program that had 30 nutrients included carotene and preformed vitamin A, as well as, total vitamin A, and included 8 additional vitamins (Thiamin, Riboflavin, Niacin, B-6, B-12, Folacin, Pathothenate, and Vitamin E) and four additional minerals (copper, magnesium, selenium and zinc). The inclusion of other items such as the subdivisions of protein and fat to animal, fish and plant fractions, were not included. In a third program that included 58 nutrients and components, there was inclusion of crude fiber, as well as, dietary fiber, linoleic and oleic fatty acids, biotin and tocopheral (as well as vitamin E), manganese and molybdenum, 11 amino acids, and water and estimated cost.

When there are missing values for specific nutrients for specific foods, they are handled either by leaving a zero or imputing the value from similar/like foods. The nutrient databases split as to how they handled this issue and how accessible the information for imputed values or missing values was made to the user.

Most of the databases were updated annually with special updates when it seemed appropriate to do so.

**Table 3 - Nutrients reported in 1976-1991 April issues of THE JOURNAL OF THE AMERICAN DIETETIC ASSOCIATION**

| 1976-1979 | 1980-1983 | 1984-1987 | 1988-1991 |
|---|---|---|---|
| Energy | Energy | Energy | Energy |
| Protein | Protein | Protein | Protein |
| | Amino acids(11) | | |
| Carbohydrate | Carbohydrate | Carbohydrate | Carbohydrate |
| | | Sugar | |
| Fat | Fat | Fat | Fat |
| | | SFA | SFA |
| | | PUFA | PUFA |
| | | | MUFA |
| | | Cholesterol | Cholesterol |
| | Dietary fiber | Dietary fiber | Dietary fiber |
| | Pectin | | |
| Vit A | Vit A | Vit A | Vit A |
| Vit C | Vit C | Vit C | Vit C |
| Thiamin | Thiamin | Thiamin | Thiamin |
| | Riboflavin | Riboflavin | Riboflavin |
| | Niacin | Niacin | Niacin |
| | Vit B-6 | Vit B-6 | Vit B-6 |
| | Vit B-12 | Vit B-12 | |
| | Folacin | Folacin | |
| | Pantothenate | | |
| Calcium | Calcium | Calcium | Calcium |
| Iron | Iron | Iron | Iron |
| | Zinc | Zinc | Zinc |
| | Magnesium | Magnesium | |
| | Phosphorus | Phosphorus | |
| | | Sodium | Sodium |
| | | Potassium | |
| | | | Chloride |
| | Alcohol | | |
| | | | Caffeine |
| n=9 | n=21 + 11 amino acids | n=24 | n=21 |

*Problem nutrients and analytic inconsistencies -*

In looking at the databases there are several problem nutrients that need to be addressed (Table 4). These problem nutrients which have incomplete data include fiber and its components, specific fatty acids, polysaccharides, disaccharides, monosaccharides, carotene, tocopherols, pantothenic acid, vitamin D, B-6, B-12, and folacin, and the minerals zinc, copper, selenium and manganese, and the various amino acids. Some data is available on all of these nutrients but not

Table 4

**Nutrients for which inadequate information exists include:**

| | |
|---|---|
| Fiber and its components | Fatty acids |
| Various amino acids | Polysaccharides |
| Disaccharides | Monosaccharides |
| Carotenes | Tocopherols |
| Pantothenic acid | Vit D |
| Vit B-6 | Vit B-12 |
| Folacin | Zinc |
| Copper | Manganese |
| Selenium | |

in all foods.

In addition to problem nutrients, there are problems with nutrient content analyses. The inconsistencies in some analytical methods make it difficult to compare data and difficult to assess its accuracy. A variation in content because of different growing conditions and different *variatils* is a strong reason for including mean plus and minus the standard deviation, or some indication of variance within the nutrient content. It is helpful when presenting data if there are some nutrients that can be included such as water value or protein to allow comparison with other analyses and other foods. It is also apparent that taking information from labels may not be an accurate way of presenting nutrient content as the label statements may vary from actual nutrient content and still be within legal boundaries.

*Needed: realistic number of significant figures -*

The number of significant figures that are presented needs to be ascertained for accuracy. With the use of computers, calculations are often carried out to an arbitrary point which is far beyond the accuracy of the data from which the value has been generated. Energy content for example, is seldom more accurate then to the calorie or a tenth of a calorie, certainly not to a hundredth, a thousandth, or ten-thousandth of a calorie. Similarly, with vitamin A values, the number of international units needs to reflect the accuracy of the original data from which those calculations were made. It is the responsibility for database providers to alert users to what a realistic number of significant figures would be and it is also the responsibility of the user to present an appropriate number of significant figures.

*Caution in using RDAs -*

RDAs, the Recommended Dietary Allowances, are extremely useful in interpreting data, however, they are an interpretation and not a reality. New editions of the RDAs change as new information is available upon which expert panels make the recommendations. Changes in the RDAs do not mean that the society has dramatic differences in the number of people who are deficient. It is critical then that one use RDA's or other kinds of recommendations with wisdom and with caution to recognize that they are interpretive. Absolute intake values should be available for accurate comparisons over time. RDAs are useful, nonetheless, in evaluating intake.

*Era-specific databases -*

It seems imperative to maintain databases that reflect specific time periods as there are changes that occur in food usage over the decades. To make useful comparisons between one era and another, one needs to have databases specific to those eras. For example, the fat content of meat has dramatically changed over the last few decades. Standards for trimming of meat have changed with a smaller amount of fat on meat after trimming. The mode of preparation changes. Genetic engineering also will result in changes within food components. Introduction of new products and cessation of the production of some products again underscores the need to maintain era-specific databases.

*Minimum specifics regarding databases to include in publication -*

There are several specifics that are desired by readers and researchers to identify databases that have been used in analysis of intake. Specifics that I feel need to be included whenever these data are presented or published are seen in Table 5. The minimum specifics include: name, year of release, version number, and a statement describing any modifications or additions. In addition, it may be appropriate to add information regarding how missing data, missing foods, or other missing data are handled, as well as, a statement as to the source or sources of the nutrient content data.

*Guidelines for selecting nutrient database programs -*

There are four guidelines that I would propose that a user should consider in selecting nutrient database programs. These steps are: 1) Analyze the needs the

Table 5

**Minimum specifics to include in publication of data generated from a nutrient database**

1. Name of database/software developer/trade name
2. Copyright year/year of most recent update
3. Version number, if applicable
4. Statement describing modifications/additions to database
5. As appropriate, a statement regarding the handling of:
   - missing data?
   - missing foods?
   - other missing data, e.g. geographic regions?

   In what detail should information regarding missing data points be given?
6. Declaration of the source(s) of nutrient data, if considered critical.

   In what detail should information regarding data source(s) be given?

user has, 2) Assess those needs with the options software programs provide, 3) Assure that the software program is updated on a regular basis, and 4) Maintain a record of the program name, version number, and latest year of the update for each specific analysis.

### Diversity of Applications and Research Needs

New research is needed to extend nutrient and food component databases (Table 6). These needs include the completion of data cells, that is the food x nutrient/ component. Analysis of additional components is also requisite. The diversity of foods and geographic diversity is not well represented in databases at this time. There is also further need for fuller inclusion of therapeutically designed foods and other engineered foods.

Table 6

**Research needed**

- Analyses to complete data cells
  - (foods x nutrients/components)
- Analyses of additional food components
  - e.g. individual fiber categories
    - individual categories of sugars
    - carotenes
    - phytochemicals
    - food additives
    - food contaminants
    - toxins, natural and contaminant
- Analyses of additional foods
  - diversity of foods
  - geographic diversity
  - therapeutically designed foods
  - engineered foods
- Estimation of the bioavailability of nutrients
- Application
  - to link nutrients/food groups/clusters
  - to assess compliance to dietary guidelines
  - to establish rational guidelines for presentation of a realistic number of significant figures individually for each nutrient/component
  - to design therapeutic diets
    - preventive diets
    - interactive-designed diets
    - designer/engineered foods
- Evaluate probability of intake, rather than solely the assumed mean intake

The bioavailability of nutrients is another logical extension for the databases to consider: Logical applications would be to cluster nutrients and food groups and to assess compliance to dietary guidelines. A further application is to develop databases that will allow the design of therapeutic diets, preventive diets, and indirective design diets. Additionally, databases can be used to design engineered foods. There also needs to be rationale developed for presenting a realistic number of significant figures for each individual nutrient or component.

As we develop databases for the 1990's, recognizing the need for excellence in diversity, it may be appropriate to evaluate the probability of intake rather than assuming a single value or mean intake. This would allow for truer representation of the society and the individual's nutrient consumption.

# Impact of Biotechnology on Future Nutrient Intakes

**Susan K. Harlander**
**University of Minnesota**
**St. Paul, Minnesota**

Biotechnology has the potential to dramatically modify the nutritional quality of the food supply. The purpose of this presentation is to provide an overview of the tools of biotechnology, how they can be used to modify plants, animals, and microorganisms used in food production, and the potential impact on the nutritional quality of the food supply.

The information contained in nutrient databases is critical to the biotechnologist. To gain regulatory approval for genetically modified foods will require comparison of engineered varieties to their traditional counterparts. If we don't have a clear understanding of the composition of traditional foods, there is no benchmark for determining in what ways the engineered counterpart has been altered or improved.

What is biotechnology? My definition of biotechnology is broad. "Bio" refers to life or living systems - if you took a biology course in high school you studied living systems - plants, animals, microorganisms and any part of these organisms. "Technology" is a method for achieving a practical purpose. Biotechnology is a collection of industrial processes that involve living systems. Biotechnology is a toolbox of techniques that can be used to modify living organisms.

Those of us in food science, have been doing "biotechnology" for thousands of years. Microorganisms such as yeast and bacteria have been used to make industrial products like bread, cheese, wine and a host of other fermented foods. Traditional plant and animal breeding and selection have also been used to improve the food supply. The original ancestors of corn look quite different than the hybrid seed corn that dominates agricultural production today. It has been modified by years of traditional breeding and selection. We have been doing biotechnology for literally thousands of years since the discovery of food fermentation and the domestication of plants and animals.

What separates "traditional" biotechnology from the "new" biotechnology is genetic engineering. Genetic engineering allows one to isolate DNA from any living organism and using restriction enzymes that function like scissors, isolate and identify specific genes. That genetic information can then be spliced or recombined with circular piecies of DNA called plasmids or vectors and transferred into a new cell (Fig. 1). DNA in all living cells is structurally and functionally identical; therefore, genetic engineering can be used to transfer DNA between organisms that do not normally mate. Human insulin can be produced by genetically engineered organisms that have received a single gene from a human pancreas cell. The bacterial cell then functions as a little factory to produce human insulin.

Genetic engineering is a powerful technology that is more precise, controllable and predictable than traditional breeding. For example, when a plant variety is crossed with a related species, all of the thousands of genes of both plants are shuffled. Genetic engineering allows the transfer of a single gene - a single trait -into a plant, an animal or a microorganism in a very predictable way because only one property is being modified. It took over 20 years to develop hybrid seed corn using traditional breeding and selection techniques; in the course of one generation, or 3 to 4 months, we can alter plants, animals, or microorganisms using genetic engineering.

Both traditional breeding and selection, and genetic engineering can be used in a number of ways to improve the food supply. Today we will focus on the use of biotechnology to enhance the nutritional quality of the foods we consume. In the immediate past, biotechnology has focussed on production agriculture

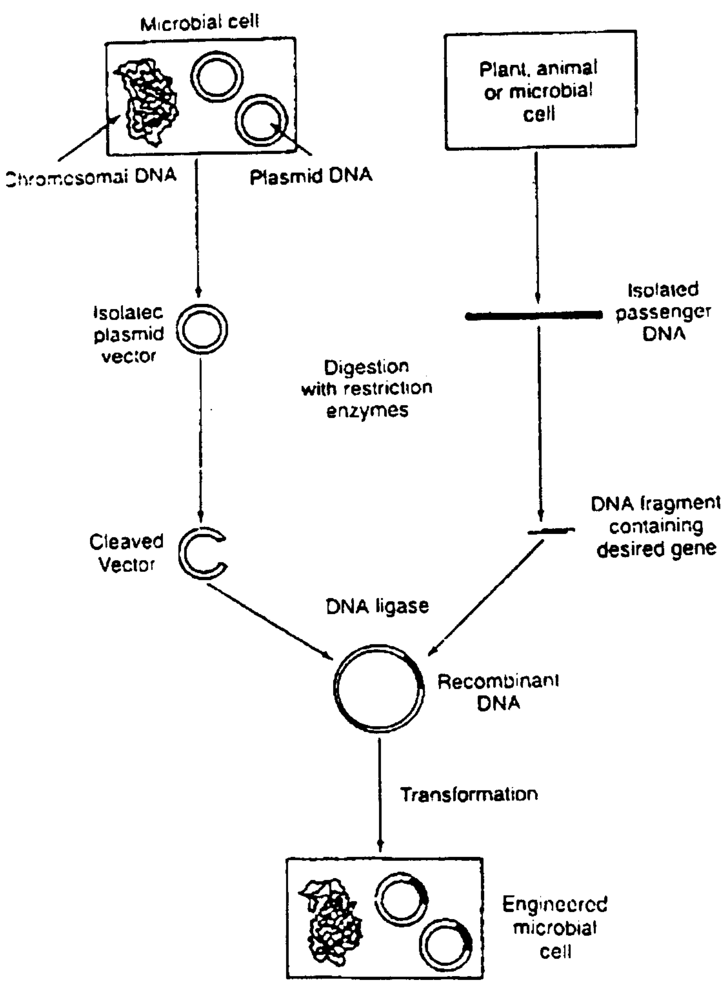



Figure 1. Gene cloning scheme in a bacterial system. A plasmid vector and passenger DNA, from plant, animal or microbial cells which codes for a specific property of interest, are isolated and subjected to restriction endonuclease digestion. The cleaved DNAs are mixed and the single stranded ends are allowed to anneal in the presence of DNA ligase, an enzyme that is capable of completing the physical joining of DNA molecules. The recombinant DNA molecules are then transformed into an appropriate host. Transformed cells are then screened for the ability to produce the gene product of interest. The appropriate cell can then be grown to produce large quantities of the gene product.

(how to produce foods more efficiently), but in the future, biotechnology will expand into value-added and consumer-driven areas.

In order to enhance the nutritional quality of the food supply, the biotechnologist needs to know which nutrients or components of foods would be desirable to increase or decrease. Interestingly, although there is consensus on the general categories of foods we should be consuming, controversy remains over the level of specific nutrients for optimum health. The 7 food groups I learned about as a child became the 4 major food groups in 1956. Even today, the debate on how to most accurately convey nutrition information to consumers continues with the proposed pyramid being considered by the U.S. Department of Agriculture. Additionally, there may be components of foods yet undiscovered that will impact nutrient requirements and health. It is probably safe to say that we do not understand everything there is to know about nutrients and nutrient interaction in our present diet. Yet this information is critical if our goal is to engineer foods with improved nutritional quality using biotechnology.

The real interest in diet right now relates to the diet/health interface. Diet has been implicated as a causative factor in a number of chronic diseases like heart disease and hypertension, osteoporosis, diabetes, obesity, cancer, aging and possibly others. Consumers are being bombarded with this information in the press. One article might discuss the concept that certain foods contain nutrients that are protective against cancer; another article indicates that the same foods contain natural toxicants that may be cancer-causing. We have overwhelmed people with information about nutrition and it's really difficult for the average consumer to sort all the information. Biotechnologists also need some help in deciding which foods and which nutrients might have the biggest nutritional impact on the largest population of consumers.

As we understand more about human genetics, we will also understand more about how an individual's diet impacts health. The human genome project will ultimately provide the genetic blueprint for humans. With that information, one should be able to design a diet to fit each individual. My nutritional needs may be quite different than yours, and these differences will be reflected in our genes. Ultimately we will have enough information to predict what the perfect diet should be, and then it is up to you and I to design that perfect diet.

I'd like to tell you about how I believe biotechnology will be used to help design the perfect diet. If one could modify the food supply using biotechnology what are some of the things that should be done to improve nutritional quality? Which nutrients should we focus on? Some that come to mind include the following:

- increasing the level of specific amino acids deficient in cereal grains
- increasing fiber content or specific components of fiber (i.e., soluble fiber)
- decreasing fat and/or cholesterol
- increasing caloric density and/or decreasing total caloric content
- increasing vitamin, mineral and trace element content
- making foods inherently digestible

Nutritionists could surely expand this list - the limit is based on one's imagination - as the possibilities are endless.

Plant biotechnologists have successfully modified the nutritional quality of protein by altering amino acid content. Cereal grains are deficient in certain amino acids and genetic engineering has been used to increase the level of these amino acids. That may not have much of an impact on nutritional quality of the food supply in this country because we have numerous protein sources in our diet. In developing countries that rely on a single cereal grain as their sole source of protein, improving the nutritional quality or the ratio of amino acids could have a major impact. The fatty acid composition of oils can also be modified. The ratio of saturated, unsaturated, and monounsaturated fatty acids in canola oil has been altered using genetic engineering. If nutritionists tell biotechnologists that it's important to increase or decrease the level of certain fatty acids in food products, genetic engineering will be able to do it.

There is concern about the safety of the food supply and most of that concern is associated with pesticide residues in foods. Pesticides are used to control insect pests. As agricultural areas interface more closely with urban areas, the concern about low level contaminants in foods will surely increase. Using genetic engineering, it is possible to build directly into plants an ability to naturally resist insects, thus decreasing the need for agricultural pesticides. A single gene from the bacterium **Bacillus thuringiensis**, an organism that has been used as a natural pesticidal agent in this country for over 20 years, can be transferred to plants. The single gene codes for a toxin that is lethal to insects, but has no effect on humans or animals, thus the engineered plant is naturally resistant to insects.

The food supply could also be modified to make it more appealing. A company in California has created the "antisense" tomato. A single gene has been removed from the tomato and reinserted back into the same plant in the opposite orientation. One sense copy of the gene the same gene inserted in the opposite direction (the `antisense' gene) blocks the action of that single gene. The particular gene that's been blocked in this example is the gene that codes for the enzyme endopolygalacturonase. This enzyme causes the tomato fruit to start breaking down when you leave it on the vine to ripen. By blocking the action of that one gene, tomato fruit can be ripened on the vine, shipped without refrigeration and has an extended shelp life. With recommendations to consume more fresh fruits and vegetables, `antisense' technology could have a major impact on the quality of fresh produce.

Many of the foods commonly consumed contain low levels of natural toxicants. This is not normally a problem as long as a well balanced diet is consumed. Overindulgence of a limited number of foods; however, could contribute to dietary carcinogen exposure. Antisense technology could be used to selectively shut off the genes involved in producing natural toxicants. Caffeine-free coffee has been developed by selectively shutting off the genes involved in caffeine production. If natural toxicants are understood, 'antisense' technology could be used to inactivate their production. In the same way, if biotechnologists understand that certain components, whether they be antioxidants, vitamins, minerals, trace elements or other compounds, are important in nutrition, then genetic engineering can be used to design foods that contain these compounds.

Antibodies are produced in humans and animals to fight diseases. Antibodies are proteins; they can be genetically engineered into plants. In the future it may be possible to deliver disease-fighting antibodies directly in the food products we consume. This may sound futuristic today, but this powerful technology can be harnessed in many ways to make the food supply healthier.

Let's say we need to construct a diet that is reduced in sodium. Can we develop things like potatoes that have natural buttery flavor and aroma, so that butter or salt doesn't need to be added? What about a variety of corn genetically engineered to produce a salty tasting protein; thus, eliminating the need for added salt? Think about the components that you perceive to be negative in the diet. There are ways to modify the food supply in such a way that the diet can deliver valuable nutrients without having to add things like butter, fat or salt. Many of the examples of plant biotechnology just discussed are currently being developed and in some cases they are already in small-scale field trials. The future is not that far away!

Biotechnology can also be applied in the animal world. Pork chops from a pig supplemented with porcine somatotropin, a growth hormone that can be produced by genetic engineering, has significantly less fat than meat from unsupplemented animals. PST causes more of the energy to be used by the pig to produce muscle rather than fat. Once it is understood how animals process nutrients into fat and cholesterol, it should be possible to genetically program cows to produce less of these compounds in meat or milk. There is tremendous concern about cholesterol in the diet. Eggs have high levels of cholesterol. Could genetic engineering be used to decrease the amount of cholesterol produced by chickens, or alter the partitioning of cholesterol in the bird? As pointed out by the previous speaker, it is critical to keep track of modifications in the food supply in nutrient databases. The food supply of the future is going to look quite different than what we have today and this historical information will help us more intelligently and expediently chart the course.

One of my areas of research focuses on the bacteria used to produce yogurt. Genetic engineering can be used to make dairy foods inherently more digestible for lactose intolerant individuals. Over 25% of the people in this country are lactose intolerant. The bacteria in yogurt help lactose intolerant individuals digest the product. These organisms produce the enzyme - galactosidase that is missing in the gut of lactose intolerant individuals. The gene for this enzyme can be transferred to bacteria used to produce other kinds of fermented dairy products.

Organisms like **Lactobacillus acidolphilus** present in Sweet Acidophilus Milk can naturally assimilate cholesterol. They are also capable of implanting in the gut, and may help your body reduce serum cholesterol levels. Other properties could also be built into bacteria and yeast using genetic engineering - whether it be anti-carcinogenic factors, stimulation of the immune response, or factors that would allow organisms to function as competitive inhibitors of other organisms present in the GI tract. The gene for the naturally sweet plant protein, thaumatin, has been inserted into yeast. Imagine making bakery products that are naturally sweet without adding sugar. As a protein, thaumatin would be digested just like any other protein. Yeast strains could be engineered to produce beta-glucans and soluble fibers that have demonstrated cholesterol-reducing properties. These could be used in a whole host of food products to help reduce serum cholesterol levels. Algae are an excellent source of beta carotine, a component implicated as an anticarcinogen. Engineered organisms could serve as sources for these valuable food ingredients.

This was a very brief overview of how biotechnology could affect the food supply - from the seed to the stomach - and beyond, to the nutritional impact of food. The goal is to deliver safe, nutritious and affordable foods that consumers want. Biotechnology is just one of many tools that will be used to provide fresh foods that are nutritious and convenient.

Todays' consumers live in urban areas - less than 2% of the population produce all the food in this country. Few of us have any understanding of how food is grown, processed, or distributed. The public is very concerned about the environment; therefore, the use of agricultural chemicals in food production is extremely important to people. Todays' consumers are also very health conscious. They read nutritional labels; they want to know what's in their food and they want to understand how that food is going to impact their health. At the same time, they are afraid of technology and they really don't have a very good understanding of the science behind biotechnology, nutrition, or food science. Science illiteracy is pervasive in this country -of 23,000 high schools in the U.S., 1900 offer no biology program, students graduating from high school are not going to have the rudamentary skills to understand what nutrition and biotechnology are all about.

Abraham Lincoln said, "Public opinion is not always right but it always prevails." We have tremendous potential to positively modify the food supply using biotechnology. If you as nutritionists will work with biotechnogists, along with biochemists, biologists, chemists, physicians, agronomists, animal scientists, and food scientists, we can create a tremendously healthful food supply. The challenge will be communicating that to the consuming public. So it's not just a matter of science and technology, it's also a matter of communication.

Activities Concerning Nutrition Labeling

# Implications for Nutrient Databases

**Janet McDonald, PhD, RD**
**US Food and Drug Administration**
**Millbrae, California**

## BACKGROUND

On March 7, 1989, Louis W. Sullivan, Secretary of the U.S. Department of Health and Human Services, calling the current food label a "Tower of Babel," announced plans for a comprehensive food labeling initiative to be undertaken by the Food and Drug Administration. In the **Federal Register** of August 8, 1989 (54 FR 32610), FDA published an advance notice of proposed rulemaking (ANPRM) that solicited public comment on a wide range of food labeling issues to help the agency determine what, if any, changes in food labeling requirements were needed to make the food label more useful and understandable to consumers. During October to December, 1989, FDA held four public hearings in Chicago, IL, San Antonio, TX, Seattle, WA, and Atlanta, GA, respectively, each focusing on a different aspect of food labeling. In the fall of 1989, in conjunction with the U.S. Department of Agriculture (USDA), FDA contracted with the National Academy of Sciences' Institute of Medicine to review current food labeling policies and recommend options for improvement. This latter effort resulted in the publication of the report entitled "Nutrition Labeling: Issues and Directions for the 1990s" (NAS-IOM, National Academy Press, Washington, D.C., 1990).

In the February 13, 1990 **Federal Register** (55 FR 5176), FDA published a reproposal of its health messages regulation, withdrawing its August 4, 1987 proposal. On July 19, 1990, FDA published in the **Federal Register** (55 FR 29476) proposed regulations on (1) Reference Daily Intakes and Daily Reference Values, (2) the mandatory status of nutrition labeling and nutrient content revision, and (3) serving sizes. These proposed rules comprised Phase 1 of FDA's food labeling initiative and incorporated comments from both the ANPRM and the four hearings. Written comments to the three proposals were solicited by the deadline of November 16, 1990. (In the same issue of the **Federal Register**, a tentative final rule on cholesterol labeling was published with a 30-day comment period.) On November 8, 1990, however, President Bush signed into law the Nutrition Labeling and Education Act of 1990 (Public Law 101-535), hereafter referred to as the NLEA or the Act, which amends the Federal Food, Drug, and Cosmetic Act "to prescribe nutrition labeling for foods, and for other purposes." The law supercedes the labeling regulations proposed by FDA. Thus, while there are many similarities between the NLEA and FDA's proposed rules, there are sufficient differences that reproposals or supplemental proposals must be made for legal and other reasons. This will necessitate the publication of more than 25 documents before the November 8, 1992 deadline for all final regulations that Congress imposed on FDA. If final regulations are not published by that date, then the proposed regulations will serve as final rules.

## NUTRITION LABELING AND EDUCATION ACT OF 1990

The major provisions of the NLEA are summarized in the appendix. Of particular note, although there are some exemptions, nutrition labeling will be required on most FDA-regulated packaged foods. Moreover, the Act calls for retailers to provide labeling information voluntarily for the 20 most frequently consumed varieties of raw fruit, vegetables, and finfish/shellfish.

In addition to the total number of calories per serving, the calories derived from total fat per serving must be declared for foods subject to mandatory labeling. The amount of the following nutrients must also be disclosed: total fat, saturated fat, cholesterol, sodium, total carbohydrates, complex carbohydrates, sugars, dietary fiber, total protein, and such vitamins, minerals, or other nutrients determined by the Secretary to be important in assisting consumers to maintain healthy dietary practices. Except for the mandatory declaration of complex carbohydrates and sugars, the NLEA contains the same nutrient labeling requirements as described in FDA's July 19, 1990 proposed regulation. The Act, however, is flexible in that it states that the Secretary may by regulation add or delete nutrients from the required list, if it is determined that such information may/may not assist consumers. Thus, it is likely that comments will be solicited regarding the declaration of complex carbohydrates and sugars. Definitions for the two terms must be established, and appropriate methods must be developed for monitoring compliance should label declaration of these components be required by the final regulation.

With regard to the mandatory declaration of saturated fat, this term also must be defined. Evidence that stearic acid does not appear to have the same cholesterol-raising propensity as other saturated fatty acids raises the question whether it should be considered a saturated fatty acid. Moreover, how should trans fatty acids be handled for labeling purposes?

Other uncertainties must be resolved with respect to dietary fiber declarations, particularly if a manufacturer wishes to highlight the content of soluble/insoluble fiber in a product. The analytical problems associated with dietary fiber measurements are discussed in the previously mentioned NAS-IOM report on nutrition labeling.

Although, heretofore, declaration of nutrients on labels has been based on analytical testing of products, in a 1979 ANPRM, FDA and USDA set forth a policy encouraging the use of properly evaluated databases for all appropriate segments of industry. Nutrient databases will constitute the means by which retailers will be expected to provide nutrition labeling information for the 20 most frequently consumed raw agricultural commodities and fish. FDA is prepared to review and evaluate analytical data generated by trade associations, or other groups, for possible approval. Presently, nutrient data are lacking for many of the several hundred species of fish and shellfish, and the extent of "natural" variation of nutrients in raw fruit, vegetables, and fish based on such factors as variety, season, source, size, and age is largely unknown.

One of the major issues with respect to the NLEA is harmonization. Not only is it important for FDA to work with USDA, which has recently announced its own labeling initiative for meat and poultry products, but we must also strive to harmonize our labeling requirements with those of Canada (with whom the United States has a free trade agreement), the European Community (which will represent a powerful economic block in 1992), and possibly Mexico (should a free trade agreement be established with that country).

## IMPLICATIONS FOR NUTRIENT DATABASES

Regarding the effect of the NLEA on databases, there is some potentially bad news, but mostly good news. The nutrients required to be declared by the NLEA will likely affect the analyses performed by or for the food industry and, ultimately, determine what goes into databases. Laboratory analyses are costly. With the proposed withdrawal of mandatory listing of thiamin, riboflavin, and niacin, for example, product analyses for these nutrients may not be performed. This could hamper dietary intake calculations not only for patients/clients but also for various research purposes, including epidemiological studies and food consumption surveys. In addition, modification of industry nutrient database systems will be necessary. Computer programs will need to reflect the new requirements for serving sizes and nutrition label content as well as for ingredient listing.

From a positive standpoint, however, because nutrition labeling will be mandatory for most foods, more analytical data will be generated, and eventually, there will be far more brand name product information available for inclusion in databases. Moreover, the new requirement for listing of such components as saturated fat, cholesterol, and dietary fiber will result in the narrowing of some data gaps. Thus, the completeness and accuracy of nutrient databases should improve because the problem of missing data or having to impute values from similar foods will be resolved to a large extent.

By November 8, 1992, FDA plans to have available a revised version of its "Compliance Procedures for Nutrition Labeling." There will be a new publication entitled "FDA Nutrition Labeling Manual: A Guide for Using Databases", which will give alternatives to current procedures.

## APPENDIX

Summary of H.R. 3562, as amended
Nutrition Labeling and Education Act of 1990

PUBLIC LAW 101-535 NOVEMBER 8, 1990

----Requires disclosure of:
- serving size in a common household measure
- # of servings per container
- # of calories per serving
- # of calories derived from total fat per serving
- Total amount of fat, saturated fat, cholesterol, sodium, sugars, dietary fiber, protein, total carbohydrates and complex carbohydrates
- Total amount of important vitamins and minerals
- *Percentage of fruit or vegetable juice contained in beverages (information panel)*
- *Mandatory ingredients in standardized foods (Effective Date 11/8/91)*
- Certified colors in ingredients list

----The Secretary may add to or delete from the above mandatory list of nutrients.

----The above labeling information would be required to be provided by retailers for the 20 most frequently consumed types of raw agricultural commodities, raw fish and shellfish if retailers fail to provide such information voluntarily. Retailers may provide this information in a single location in their store. Effective date for voluntary labeling: 11-8-91.

----Food exempted from the labeling requirement:
- Food sold in restaurants
- Food sold at prepared food counters in grocery stores
- Infant formula
- Medical foods
- Food sold in bulk between wholesalers
- Food sold in small packages have a partial exemption
- Food with insignificant amounts of nutrients
- Food sold by merchants with total sales of not more than $500,000/year or food sales not more than $50,000/year

----Except as noted above, regulations to implement these nutrition labeling requirements must be promulgated within 24 months and must, among other things,
- permit the voluntary addition of nutrition infor mation of the same type that is required by law.
- permit nutrition information and labeling to remain the same on similar foods or foods packed as an assortment, even though there are minor nutritional variations among such foods.
- require the nutrition information to be presented in a manner which enables consumers to under stand the relative significance of the information in the context of a total daily diet.

----Removes food standards (except those for dairy products and maple syrup) from formal rulemaking requirements.

----Nutrient Content Claims
- A claim which characterizes the level of one or more of the required nutrients (such as high fiber, low sodium, or no cholesterol) is allowed only if the characterization of the level uses terms which are defined by the Secretary.
- No claim with respect to cholesterol may be made if the food contains an amount of fat or saturated fat that increases the risk of a health-related condition or disease, except under certain limited circumstances.
- No claim with respect to saturated fat may be made if the food contains cholesterol, unless the amount of cholesterol is disclosed with the claim.
- No claim with respect to dietary fiber may be made, unless the food is low in total fat or the level of total fat is disclosed in the claim.
- For all nutrient content claims, a statement direct ing the consumer to the nutrition information panel must appear in appropriate proximity to such claim, and if the product contains any other nutrients in amounts that increase the risk of a health-related condition or disease, the state ment must identify the nutrient.
- The Secretary must define "free", "low", "light/lite", "reduced", "less", and "high".
- Regulations must be promulgated in 24 months.

----Health Claims
- A claim which characterizes the relationship of one or more of the required nutrients to disease (such as fiber preventing cancer, or low choles terol preventing heart disease) may only be allowed if such claim is authorized by the Sec retary as scientifically valid, and if the other

required nutrients are present in amounts which do not increas the risk of a health-related condition or disease (unless the Secretary finds that the claim would assist consumers).

- Claims involving voluntarily labeled nutrients must be made in the same manner as with claims for nutrients required to be on the label.
- The Secretary will determine the appropriate standard and procedure for claims made regard -ing vitamins, minerals and other dietary supple -ments.
- Regulations must be promulgated in 24 months.

----State Enforcement

- States are authorized to enforce these new label -ing requirements on behalf of the Federal Government (30 days advance notice required).

----National Uniformity

- The bill provides for national uniformity in nutri -tion labeling and claims and in certain other labeling matters.

----Implementation Date

- Except as noted, labeling regulations must be implemented by May 8, 1993.

# Nutrition Labeling: Problems or Opportunities?

**Nancy J. Tucker, CAE**
**Produce Marketing Association**
**Newark, Delaware**

The sub-title of my speech on nutrition labeling is "Problems or Opportunities"? These two words are not necessarily an either/or scenario. Throughout history, innovative people have considered problems as challenges and made them into exciting opportunities.

I'm here to talk about the problems and resulting opportunities with nutrition labeling for fresh produce. Although such problems and opportunities have been around for the past decade, things have really heated up since The Food and Drug Administration's proposed regulations and the passage of the Nutrition Labeling and Education Act of 1990.

To provide background and foundation, I'd like to talk briefly about some of the past problems with nutrition labeling and how these became opportunities. Then I'll focus on the impact of the legislation and FDA's corresponding regulations.

### [1]Problem #1: What data base to use?

In the late 1970"s, the produce industry wanted to respond to consumer's budding interest in nutrition by promoting the healthful benefits of fresh fruits and vegetables. Even though fresh produce was exempt from nutrition labeling, voluntary efforts needed FDA's approval. Unfortunately, FDA did not approve of using existing data bases (such as USDA's Handbook .8) for labeling fresh produce.

### [2]Opportunity #1: Provide a data base for the industry.

Working with FDA, and using an approach pioneered by The Potato Board, PMA began nutrition research for fresh fruits and vegetables. At first, only a few commodities were studied. Then, year by year, momentum grew as more organizations saw the value of this grower funded research.

Here's how we carried out the research. [3]Produce samples were collected from more than 16 cities across the country at different times during the growing season to take into account the many variables of seasonality, growing conditions, variety and handling. As a result, the nutrition label for a potato could be used for a red, white or russet. Consumers were assured, within a 95% confidence level, that the nutrients on the label were in the potato they bought.

[4]By the end of this year, research will be complete on 42 commodities. The values from the research have been accepted by FDA for nutrition labeling, or in the case of more recent studies, are still under review.

I'm going to jump ahead here to a corollary problem created by the new legislation. The legislation said FDA must provide the values for the nutrition labels of the top 20 fruits and top 20 vegetables. PMA has conducted research for most, but not all of these commodities. (Some of the items we've researched are not among the top 40.) FDA will not spend their already tight resources and staff to conduct nutrition research on the remaining commodities.

USDA data will be used for the remaining commodities. However, USDA goes no further than providing a mean. It does not take into account variability or a statistical confidence level. FDA is wrestling with this issue now. One of PMA's members, Data Documents Systems completed these extensive computations using USDA data and have submitted them to FDA.

### [5]Problem #2: How to effectively communicate nutrition information to consumers?

A nutrition label by itself is of limited use to consumers or to marketers trying to promote produce.

**[6]Opportunity #2: Create a marketing tool.**

Retailers and their advertising departments needed help in communicating and marketing nutrition information. [7]Therefore, PMA developed a collection of marketing tools, the Nutrition Marketing Resource for Produce. The book contains [8]camera-ready POP signs, [9]advertising slogans that conform to FDA's current guidelines, [10]nutrient charts, [11]healthy recipes and more. Marketing pages on newly researched commodities and updates on government regulations have been provided at no extra charge.

[12]The Nutrition Marketing Resource was first published three years ago. Using the tools in this book, many retailers started nutrition information programs. Point of sale sign companies used the information in their offerings.

**New Legislation**

[13]Enough of history. Let's move on to last summer's events which had a profound impact on the industry's nutrition labeling efforts. A new set of problems, and opportunities, was created.

In July, the FDA proposed comprehensive regulations for nutrition labeling of all foods, including mandatory labeling for fresh fruits and vegetables. In the fall, the Nutrition Labeling and Education Act of 1990 was passed.

The legislation provided the outline for *voluntary* labeling of the 20 most frequently consumed fruits and 20 most frequently consumed vegetables. However, retailers must substantially comply with these voluntary measures or they become *mandatory*. This brings us to the next problem.

**[14]Problem #3: What are the top 20 fruits and top 20 vegetables?**

How do you determine the 20 most frequently consumed fruits and 20 most consumed vegetables? One way is to measure dollar amount sold. But this data is hard to obtain and it emphasizes the more expensive items. How about measuring sales by volume (number of pounds sold)? This may skew the results in favor of heavier items. Another alternative is to obtain data from consumption survey? Unfortunately, consumption data tends to be old and may be limited in terms of the number of commodities it covers.

**[15]Opportunity #3: Conduct research and provide information to FDA.**

PMA felt it was very important to provide input into the process of determining the top 40 items. We chose to conduct original research last fall on the amount of produce sold by volume, judging this was the most accurate and feasible of the three approaches. Retailers were asked to report 12 month sales figures, by volume, for their top 25 commodities. The data was compiled and submitted to FDA. FDA staff were pleased to receive PMA's list of the top 20 fruits and top 20 vegetables, and were happy to see that it was very similar to information they compiled from consumption studies.

**[16]Problem #4: How to standardize serving sizes for fresh produce?**

Another hurdle posed by FDA's regulations is that of establishing standardized serving sizes for fresh fruits and vegetables. Probably no food in the supermarket has more sizes, shapes, and varieties than fresh produce. Yet it is important to have standardized serving sizes to provide consistency and allow consumers to make comparisons.

**[17]Opportunity #4: Serving size ranges.**

FDA developed a good system of serving size ranges for fresh fruits in last summer's nutrition labeling proposal. A centralized serving size of 5 oz. was established. Items that fell into the range of 50% above or below that central figure could be labeled with an average weight. An example of this would be an [18]orange which could be labeled at its average weight of 6 oz. Commodities such as [19]strawberries, which fall below the range or [20]cantaloupe, which fall above the range, would be labeled as 5 oz.

Good data on the "average" size for a fruit or vegetable comes from PMA's nutrition research. [21]During the sampling process at least one hundred and sometimes many hundreds of individual items were purchased from supermarket shelves and sent to the laboratory where they were weighed and measured. This data supports the 5 oz. serving size suggested by FDA. It also supports extending the same range to fresh vegetables. Extensive comments on serving sizes for fruits and vegetables were developed by both PMA and United Fresh Fruit and Vegetable Association and submitted to FDA.

**[22]Problem #5: Displaying nutrition information in the stores.**

The legislation states that nutrition labels of the top 40 items must be displayed in the area where the commodities are offered for sale. This presents a logistical problem. Do you use consumer brochures? If so, then sufficient quantities must be printed, stocked, displayed and maintained. Do you use a poster? If may be difficult to hang a poster in produce departments which have little or no free wall space. Do you use a reference book? Again, where do you put it, and, how do you keep it from walking? Even if chained down, pages of a book can still be ripped out. Computers (information kiosks) could be a good solution to the problem, but they have a high capital cost.

Another concern related to displaying nutrition information is the uncertainty of what the new label will look like. As far as we can gather from FDA, their final recommendations for a new label format will not be published until 1992. Waiting until then to develop nutrition materials for fresh produce would greatly delay implementation of labeling programs and make it hard for the industry to substantially comply with the legislation by the existing deadlines. A temporary solution to this problem may be presented in FDA's publication of "phase one" of their nutrition labeling regulation reproposal which is expected in the Federal Register any day now.

**[23]Opportunity #5: Develop materials for retailers.**

The opportunity for retailers and suppliers to develop nutrition marketing programs is wide open. Now is the time innovative companies should be working to meet the needs of retailers.

Unfortunately, not enough research has been conducted on what consumers judge to be the most effective way to provide nutrition information. PMA just scratched the surface with a small study conducted last year with more than 200 shoppers in northern California. [24]And what did the shoppers say was the most useful way to provide this information? Signs by the produce items were most often cited (67%), with take-home brochures coming in second with 57% of the vote.

As long as they conform to FDA's regulation, any of these materials will help retails comply with the legislation.

**[25]Problem #6: What is substantial compliance?**

In May, 1993, the legislation requires FDA to produce a report stating whether industry has substantially complied with voluntary nutrition labeling. If substantial compliance is not found, the voluntary regulations become mandatory. The problems here are how to define substantial compliance, how to measure it and when to measure.

**[26]Opportunity #6: Work with FDA on these issues.**

Trade associations and industry members need to present comments to FDA on defining and measuring substantial compliance.

It is also the responsibility of groups like PMA to encourage retailers to comply with the voluntary labeling regulations.

**[27]Problem #7: How can we encourage retailers to comply with the voluntary regulations?**

Fortunately, this does not look like a hard job because many retailers are chomping at the bit to implement nutrition labeling programs

However, while some retailers may be anxious to start a labeling program, other may need more convincing.

**[28]Opportunity #7: Sell the benefits of nutrition labeling.**

We need to sell the benefits of nutrition labeling. We need to convince retailers to do it, not just because the law says so, but because it is in their own best interests!

Three key benefits of nutrition labeling are it's value as a customer service tool, as a way to increase customer perception of the store and it's use in advertising.

**Customer Service Tool**

[29]The use of nutrition information in the produce department is an important customer service tool. The number of consumers who rate the availability of nutrition and health information as a very or somewhat important factor when they select a supermarket has been steadily increasing over the past 4 years.

Consumer interest in nutrition will continue to grow as a function of demographics. People become more interested in nutrition as they get older. [30]These kids may be fairly typical, as they snack on chips and soda and say "It's a good thing we're not hung up on nutrition yet!"

[31]However, our population as a whole is getting older. People age 65 and older traditionally have been most concerned about the nutritional content of the foods they eat. They also are the fastest growing population group. The 50-64 year olds have the highest average income per household. Many of the baby boomers are going into their forties - and their interest in nutrition continues to grow, as shown in this slide.

[32]Consumers want nutrition information, they want it in the produce department, and they have a higher

opinion of store who provide it as opposed to those who don't.

### Increased consumer perception for store

[33]A study on the impact of nutrition information in the produce department was conducted by researchers at the University of Santa Clara, California and published in the Journal of Retailing in 1987. Six items in the produce department were highlighted with nutrition signs. Even though the number of signed items was small, the researchers found that consumers had a significantly increased perception of the supermarkets with the nutrition signs. This was compared to control stores using POP signs with no nutrition information (just buying and storing tips) and stores with signs stating just price.

[34]In addition, the researchers made an unexpected and interesting finding. Consumers thought a fruit or vegetable with a nutrition sign on it was of higher quality than a produce item that did not have a nutrition sign. They did not necessarily think it more nutritious, just of higher quality. Our best guess is that the nutrition sign acted as a type of "seal of approval".

### Valuable advertising/sales tool

[35]This brings me to the third benefit of nutrition marketing - advertising and sales. Encouraging more purchases of fruits and vegetables is the main goal of any produce nutrition marketing program.

[36]The Point of Purchase Institute says that consumers make 80% of their overall food purchasing decisions at the grocery store shelf. The percentage for produce is even greater since it is an especially high impulse purchase. Since 96% of consumers say that nutrition is a very or somewhat important factor when they purchase food, a short, interesting nutrition message could make the difference between a sale or no sale.

It has been extremely difficult to measure the sales impact of nutrition signs and advertising since so many variables affect the sales of fresh produce. However, a couple recent studies have shown a small increase in sales due to a nutrition marketing program.

### [37]Problem #8: How do you effectively communicate the nutrition message?

The real goal of the legislation is to give consumers more information so they can make informed food choices and include more healthy foods in their diet. While a label may be helpful to consumers seeking such information, it may not make an impact on many Americans.

### [38]Opportunity #8: Programs and research.

We will see a large number of new educational and marketing programs in the next 18 months. For example the National Food Processors Association is spearheading the Food Labeling Education Project which plans to produce a guide for professionals to help consumers understand and use the new food label to make informed food choices.

The Food and Drug Administration will also take an active role in educating consumers. The nutrition labeling legislation gives FDA the authority to develop consumer education programs. This is quite noteworthy since it is one of the few (if not the only) mandates FDA has had to develop consumer programs. After the final label format is developed, FDA plans to do research on what consumer messages will be needed and effective.

### Implementing nutrition program in department

Since some of you may be planning or developing nutrition labeling programs, I thought I's pass along some tips gleaned from research and a review of successful nutrition programs:

[39]* *Develop messages relevant to consumer interests and concerns.* The study in northern California showed stores had successful programs when highlighting one particular nutrition story, in this case, fiber. A rotating schedule of promotions centered on individual nutrients such as fiber, calories, vitamin A, potassium or others could provide an exciting promotional schedule.

[40]* *Use point-of-purchase shelf labels as cues.* [41]When consumers in the California study ranked the most useful ways to get information, shelf signs were at the top of the list. The impulse nature of produce buying decisions makes it a natural for nutrition messages on shelf labels. But the message needs to be short and simple. A complete nutrition label over each produce item is unlikely to make much of an impact on the consumer.

[42]* *Use sources of information that are credible.* Fortunately, we will have a uniform nutrition values, provided by FDA, compiled from PMA data and USDA data.

[43]* *Design highly visible materials distinguished from commercial promotions.* It's not easy to cut through the clutter of materials and information overload that consumers are exposed to. Again, I stress that nutrition marketing signs and information need to be simple and concise.

[44]* *Make it long term.* Nutrition promotions are a long-term, consumer education service. They have some of

their greatest impact on consumer perceptions and loyalty.

[45] * *Coordinate the program with supermarket personnel.* If the program is effective, it will spark consumer interest and questions. The people receiving many of these questions will be the produce clerks and managers. They need the resources to answer these questions.

And last but not least is the suggestion to coordinate efforts with those of health organizations and other associations.

For example, the new dietary guidelines urge Americans to eat 5 or more servings of fresh fruits and vegetables each day. [46]This recommendation forms the focus of the new Produce for Better Health Foundation. The purpose of this organization is to help alleviate certain chronic disease conditions by encouraging greater fruit and vegetable consumption and by awarding grants to institutions to help with this effort. The first board meeting of the organization is next week, but plans are already in the works to have retail materials ready by mid-fall.

The industry can also work with health organizations such as the American Heart Association which sponsors a healthy Food Festival each fall and the [47]American Cancer Society whose last two April campaigns have focused on healthy foods such as fresh fruits and vegetables to reduce the risk of cancer.

**[48]Win - Win situation**

Using nutrition marketing in the produce department is a win-win situation. Consumers win as they learn how fruits and vegetables can help them live healthier lives. The produce industry wins as people buy more produce.

You see, nutrition labeling is not a problem - it's an exciting challenge and a wonderful opportunity.[49]

**Endnotes**

1 Problem #1: What data base to use?
2 Opportunity #1: Provide a data base for the industry.
3 CN in Supermarket
4 Nutritional Label
5 Problem #2: How to effectively communication nutrition information to consumers?
6 Opportunity #2: Create a marketing tool.
7 Nutrition Marketing Resource Book
8 POP sign.
9 Advertising slogans.
10 Nutrient chart.
11 Healthy recipe.
12 Nutrition Marketing Resource Book picture.
13 Blank slide.
14 Problem #3: What are the top 20 fruits and top 20 vegetables?
15 Opportunity #3: Conduct research and provide infor -mation to FDA.
16 Problem #4: How to standardize serving sizes for fresh produce?
17 Opportunity #4: Serving size ranges.
18 Orange picture.
19 Strawberry picture.
20 Cantaloupe picture.
21 CN in supermarket.
22 Problem #5: Displaying nutrition information in the stores.
23 Opportunity #5: Develop materials for retailers.
24 Best ways to provide information in stores.
25 Problem #6: What is substantial compliance?
26 Opportunity #6: Work with FDA on these issues.
27 Problem #7: How can we encourage retailers to comply with the voluntary regulations?
28 Opportunity #7: Sell the benefits of nutrition labeling.
29 Customer service competitive tool
30 Kid's Cartoon
31 Customer concern
32 Produce Department shot
33 Consumers had significantly higher perception w/ nutrition POP
34 Perception of higher quality
35 Effective Advertising and Sales Tool
36 Decisions made at supermarket shelf
37 Problem #8: How do you effectively communicate the nutrition message?
38 Opportunity #8: Programs and research.
39 Successful nutrition programs
40 Successful nutrition programs
41 Most useful way to provide information
42 Successful nutrition programs
43 Successful nutrition programs
44 Successful nutrition programs
45 Successful nutrition programs
46 5 - a - Day slide.
47 ACS program in action.
48 Nutrition Marketing-A Win-Win Situation
49 PMA Logo

Industry Response to Labeling Initiatives: Impact on Industry Activities Regarding Food Analysis and Nutrient Databases

# "A Perspective from the Baby Food Industry"

**Sandra J. Bartholmey**
**Gerber Food Products**
**Fremont, Michigan**

## INTRODUCTION

Baby food manufacturers have provided mandatory nutrition information on baby food labels since the first nutrition labeling guidelines were approved in 1973 for products that are regulated by FDA and since 1983 for USDA regulated products which contain meat and poultry. Although approximately thirty percent of food companies have been providing nutrition information on a voluntary basis since these regulations were made, the 1990 proposed rules published by FDA, the Nutrition Labeling and Education Act of 1990 and USDA's Advance Notice of Proposed Rulemaking on Nutrition Labeling of Meat and Poultry Products published this April are the first concerted efforts to mandate nutrition labeling of almost all foods.

For food companies such as Gerber that already have nutrition labeling programs in place, the impact of the new regulations is different from that on companies that do not have such programs in place. However, I want to emphasize the profound impact these regulations will have, even on a company with analytical capabilities and data base in place that satisfies current nutrition labeling mandates.

But first, I want to quickly review the rationale that inspired the new regulations and look at it from a baby's perspective because I think this is an important point we need to keep in mind in our discussions.

## RATIONALE FOR NUTRITION LABELING CHANGES

The 1988 **Surgeon General's Report on Nutrition and Health** presented the scientific basis that related dietary excess and imbalance to the major causes of death and chronic disease in the United States. In 1989, the National Research Council's Committee on Diet and Health issued its Report on **Diet and Health: Implications for Reducing Chronic Disease Risk**. This report also focused on clarifying "the role of diet in the etiology and prevention of the major causes of morbidity and mortality in the United States."

Both reports recommended dietary modifications along the lines of the **Dietary Guidelines for Americans** to help the majority of Americans become healthier. These recommended dietary changes for adults are the driving force behind the nutrition labeling changes designed to help consumers reduce intake of foods high in fat and saturated fat and increase intake of foods high in complex carbohydrate and fiber. These recommendations are sound advice for most adults and older children.

## BABIES ARE DIFFERENT FROM ADULTS

Both reports, however, recognize that the dietary needs of infants and young children are different from the recommendations made for adults to modify their diets, that the nutritional priorities of infants and young children for growth and development require different dietary recommendations.

Gerber thinks the labeling of foods should reflect that difference by exempting baby foods from labeling of saturated fats, cholesterol and fiber -- nutrients that have a reverse importance in the infant diet compared to the adult diet. This position will be discussed during the panel discussion.

## NUTRITION LABELING CHANGES

I want to give credit to several people who provided the backbone of this presentation. First of all to Bonita Funk, Manager of Gerber's Nutrition Lab, who is the meticulously thorough person who compiled the following information for a presentation to the Foods Lab Conference held in Pittsburgh last month on the Impact of Labeling Changes on the Food Lab. Jim Kraley helped with information on label design, and Wesley Meeuwsen on nutrient data banks.

The nutrients currently required on labels are listed on the left of the slide, the FDA proposed nutrients in the center and those required by NLEA on the right. As can be seen from the highlighted areas, both of the new formats differ significantly from the current one.

Declarations of percent U.S.RDA for protein, thiamin, riboflavin, and niacin have been eliminated. Declarationsof percent of Reference Daily Intake (RDI) for Vitamin A, Vitamin C, Calcium and Iron are required.

A calculated value for calories from fat will now be required.

Declarations of saturated fatty acids and cholesterol will be required.

NLEA of 1990 requires carbohydrates to be further categorized as complex carbohydrates, sugars and dietary fiber. The FDA proposes an additional declaration of dietary fiber.

In summary, the changes call for additional labeling of calories from fat, saturated fatty acids, cholesterol, sugars and dietary fiber. How will these affect analytical capabilities, and what are other considerations regarding changes in labels?

## IMPACT ON GERBER FOOD ANALYSES METHOD DEVELOPMENT

*Fatty Acids*

Gas chromatography methods now in place can adequately quantify saturated fatty acids. Better methods need to be developed for mono- and polyunsaturates. Gerber's Nutrition Lab is participating with other food companies in a collaborative study to determine which methods and GC columns give the best resolution of the unsaturated fatty acids in various food products. Once an updated method is established, we can verify our existing fatty acid data or start re-analyzing our products. Collaborative studies are time-consuming, but necessary to satisfy AOAC requirements for validated methodology.

*Cholesterol*

An AOAC-approved cholesterol method is in place, but is time-consuming, tedious and requires the use of benzene, a potent liver carcinogen. Again, a collaborative study is needed to validate an alternative method that was developed that appears to be accurate, saves time and does not require benzene.

*Sugars*

High Performance Liquid Chromatography (HPLC) is the method of choice for sugar analysis. FDA proposes to use HPLC to monitor label compliance for sugar content and proposes to include tri- and tetrasaccharides and sugar alcohols in with the more commonly defined sugar content of mono- and disaccharides. This presents a challenge for food labs because current HPLC methods do not all determine sugar alcohols, and no collaborative studies have validated methods that determine saccharides beyond two glucose units.

*Dietary Fiber*

Analysis of dietary fiber by the AOAC-approved method (aka Prosky method) is the method specified for use in FDA's proposed ruling. The Prosky method requires three days and extensive analyst time to perform. Two less labor-intensive methods have been reviewed by collaborative study, but more validation is required before one can be adopted for official use.

In summary, better analytical methods that are more efficient and at least as accurate at current AOAC methods are needed to help provide accurate information to consumers at a reasonable cost.

## COSTS

*Laboratory Analyses*

Speaking of costs, the nutrition labeling changes will cost Gerber anywhere from $90,000 to $172,000 in analytical costs alone based on an estimated $500 to $960 per product. These estimates are based on analysis of one sample from each of three packs where a sample is defined by the FDA sampling protocol as a composite of twelve containers pulled at random throughout a pack period. We consider data from three "samples" the minimum on which to base a label declaration, following USDA policy.

The Gerber in-house cost estimates shown in this slide are based on our current average hourly laboratory charge rate and do not include reagents or other supplies, equipment or additional personnel that may be required for the analyses. Because Gerber has an in-house lab and equipment and chemists in place, the lab costs look favorable compared to the higher cost of a contract lab. But for companies without such facilities, the contract lab may be the least expensive option in the long run.

*Changes in Label Format Design*

With each new line of nutrient data added to a label an inch wide, we challenge our Corporate Design department with a major undertaking. Once they have managed to squeeze another item in the limited space allotted and the Legal, Nutrition, Label Control, and Marketing departments have reviewed it and approved it, then USDA has to approve it if it is a meat or poultry item. Finally, Purchasing can order the label to be printed and arrive to coincide with the production schedules in the plants.

The average cost for the design changes we have been discussing are anticipated to be about $1200 per label which totals $216,000 for 180 products in the Gerber line of baby foods.

*Nutrient Data Bank*

Adding more nutrients to our nutrient data bank does not present a problem. Slots are available to receive added nutrients for each product, and programs are in place to calculate raw data into amounts of nutrients per serving and per 100 grams. The biggest challenge will be to generate the data and make whatever changes will be required on the label within the time frame proposed for implementation of the nutrition labeling changes, by May, 1993.

## SUMMARY

The impact of nutrition labeling changes on Gerber Products Company can be summarized in economic terms. Total costs for nutrition labeling changes are estimated to fall between $300,000 and $400,000. Since the analyses must be repeated annually to verify label accuracy, this range represents only the initial impact. Another component of the NLEA of 1990 is the educational element. Whether or not baby foods will be required to carry all of the new nutrient information, it will be essential to write new or revise current consumer and professional publications to include the new nutrient data and place them in proper perspective regarding their importance in the infant diet.

## BIBLIOGRAPHY

Ellefson, Wayne; (1991) Nutrition labeling--progressing to meet consumer needs. Hazleton Food Science Newsletter, 2 (38):1.

Nutrition Policy Board, USDHHS; (1988) The Surgeon General's Report on Nutrition and Health. U.S. Department of Health and Human Services, Public Health Service, DHHS (PHS) Publication No. 88-50210. U.S. Government Printing Office, Washington, DC 20402.

Committee on Diet and Health, Food and Nutrition Board; (1989) Diet and Health: Implications for Reducing Chronic Disease Risk. Commission on Life Sciences, National Research Council, National Academy of Sciences, National Academy Press, Washington, D.C.

Industry Response to Labeling Initiatives: Impact on Industry Activities Regarding Food Analysis and Nutrient Databases

# "A Perspective from the Baby Food Industry"

**Sandra J. Bartholmey**
**Gerber Food Products**
**Fremont, Michigan**

## INTRODUCTION

Baby food manufacturers have provided mandatory nutrition information on baby food labels since the first nutrition labeling guidelines were approved in 1973 for products that are regulated by FDA and since 1983 for USDA regulated products which contain meat and poultry. Although approximately thirty percent of food companies have been providing nutrition information on a voluntary basis since these regulations were made, the 1990 proposed rules published by FDA, the Nutrition Labeling and Education Act of 1990 and USDA's Advance Notice of Proposed Rulemaking on Nutrition Labeling of Meat and Poultry Products published this April are the first concerted efforts to mandate nutrition labeling of almost all foods.

For food companies such as Gerber that already have nutrition labeling programs in place, the impact of the new regulations is different from that on companies that do not have such programs in place. However, I want to emphasize the profound impact these regulations will have, even on a company with analytical capabilities and data base in place that satisfies current nutrition labeling mandates.

But first, I want to quickly review the rationale that inspired the new regulations and look at it from a baby's perspective because I think this is an important point we need to keep in mind in our discussions.

## RATIONALE FOR NUTRITION LABELING CHANGES

The 1988 **Surgeon General's Report on Nutrition and Health** presented the scientific basis that related dietary excess and imbalance to the major causes of death and chronic disease in the United States. In 1989, the National Research Council's Committee on Diet and Health issued its Report on **Diet and Health: Implications for Reducing Chronic Disease Risk**. This report also focused on clarifying "the role of diet in the etiology and prevention of the major causes of morbidity and mortality in the United States."

Both reports recommended dietary modifications along the lines of the **Dietary Guidelines for Americans** to help the majority of Americans become healthier. These recommended dietary changes for adults are the driving force behind the nutrition labeling changes designed to help consumers reduce intake of foods high in fat and saturated fat and increase intake of foods high in complex carbohydrate and fiber. These recommendations are sound advice for most adults and older children.

## BABIES ARE DIFFERENT FROM ADULTS

Both reports, however, recognize that the dietary needs of infants and young children are different from the recommendations made for adults to modify their diets, that the nutritional priorities of infants and young children for growth and development require different dietary recommendations.

Gerber thinks the labeling of foods should reflect that difference by exempting baby foods from labeling of saturated fats, cholesterol and fiber -- nutrients that have a reverse importance in the infant diet compared to the adult diet. This position will be discussed during the panel discussion.

## NUTRITION LABELING CHANGES

I want to give credit to several people who provided the backbone of this presentation. First of all to Bonita Funk, Manager of Gerber's Nutrition Lab, who is the meticulously thorough person who compiled the following information for a presentation to the Foods Lab Conference held in Pittsburgh last month on the Impact of Labeling Changes on the Food Lab. Jim Kraley helped with information on label design, and Wesley Meeuwsen on nutrient data banks.

The nutrients currently required on labels are listed on the left of the slide, the FDA proposed nutrients in the center and those required by NLEA on the right. As can be seen from the highlighted areas, both of the new formats differ significantly from the current one.

Declarations of percent U.S.RDA for protein, thiamin, riboflavin, and niacin have been eliminated. Declarationsof percent of Reference Daily Intake (RDI) for Vitamin A, Vitamin C, Calcium and Iron are required.

A calculated value for calories from fat will now be required.

Declarations of saturated fatty acids and cholesterol will be required.

NLEA of 1990 requires carbohydrates to be further categorized as complex carbohydrates, sugars and dietary fiber. The FDA proposes an additional declaration of dietary fiber.

In summary, the changes call for additional labeling of calories from fat, saturated fatty acids, cholesterol, sugars and dietary fiber. How will these affect analytical capabilities, and what are other considerations regarding changes in labels?

## IMPACT ON GERBER FOOD ANALYSES METHOD DEVELOPMENT

*Fatty Acids*

Gas chromatography methods now in place can adequately quantify saturated fatty acids. Better methods need to be developed for mono- and polyunsaturates. Gerber's Nutrition Lab is participating with other food companies in a collaborative study to determine which methods and GC columns give the best resolution of the unsaturated fatty acids in various food products. Once an updated method is established, we can verify our existing fatty acid data or start re-analyzing our products. Collaborative studies are time-consuming, but necessary to satisfy AOAC requirements for validated methodology.

*Cholesterol*

An AOAC-approved cholesterol method is in place, but is time-consuming, tedious and requires the use of benzene, a potent liver carcinogen. Again, a collaborative study is needed to validate an alternative method that was developed that appears to be accurate, saves time and does not require benzene.

*Sugars*

High Performance Liquid Chromatography (HPLC) is the method of choice for sugar analysis. FDA proposes to use HPLC to monitor label compliance for sugar content and proposes to include tri- and tetrasaccharides and sugar alcohols in with the more commonly defined sugar content of mono- and disaccharides. This presents a challenge for food labs because current HPLC methods do not all determine sugar alcohols, and no collaborative studies have validated methods that determine saccharides beyond two glucose units.

*Dietary Fiber*

Analysis of dietary fiber by the AOAC-approved method (aka Prosky method) is the method specified for use in FDA's proposed ruling. The Prosky method requires three days and extensive analyst time to perform. Two less labor-intensive methods have been reviewed by collaborative study, but more validation is required before one can be adopted for official use.

In summary, better analytical methods that are more efficient and at least as accurate at current AOAC methods are needed to help provide accurate information to consumers at a reasonable cost.

## COSTS

*Laboratory Analyses*

Speaking of costs, the nutrition labeling changes will cost Gerber anywhere from $90,000 to $172,000 in analytical costs alone based on an estimated $500 to $960 per product. These estimates are based on analysis of one sample from each of three packs where a sample is defined by the FDA sampling protocol as a composite of twelve containers pulled at random throughout a pack period. We consider data from three "samples" the minimum on which to base a label declaration, following USDA policy.

The Gerber in-house cost estimates shown in this slide are based on our current average hourly laboratory charge rate and do not include reagents or other supplies, equipment or additional personnel that may be required for the analyses. Because Gerber has an in-house lab and equipment and chemists in place, the lab costs look favorable compared to the higher cost of a contract lab. But for companies without such facilities, the contract lab may be the least expensive option in the long run.

*Changes in Label Format Design*

With each new line of nutrient data added to a label an inch wide, we challenge our Corporate Design department with a major undertaking. Once they have managed to squeeze another item in the limited space allotted and the Legal, Nutrition, Label Control, and Marketing departments have reviewed it and approved it, then USDA has to approve it if it is a meat or poultry item. Finally, Purchasing can order the label to be printed and arrive to coincide with the production schedules in the plants.

The average cost for the design changes we have been discussing are anticipated to be about $1200 per label which totals $216,000 for 180 products in the Gerber line of baby foods.

*Nutrient Data Bank*

Adding more nutrients to our nutrient data bank does not present a problem. Slots are available to receive added nutrients for each product, and programs are in place to calculate raw data into amounts of nutrients per serving and per 100 grams. The biggest challenge will be to generate the data and make whatever changes will be required on the label within the time frame proposed for implementation of the nutrition labeling changes, by May, 1993.

## SUMMARY

The impact of nutrition labeling changes on Gerber Products Company can be summarized in economic terms. Total costs for nutrition labeling changes are estimated to fall between $300,000 and $400,000. Since the analyses must be repeated annually to verify label accuracy, this range represents only the initial impact. Another component of the NLEA of 1990 is the educational element. Whether or not baby foods will be required to carry all of the new nutrient information, it will be essential to write new or revise current consumer and professional publications to include the new nutrient data and place them in proper perspective regarding their importance in the infant diet.

## BIBLIOGRAPHY

Ellefson, Wayne; (1991) Nutrition labeling--progressing to meet consumer needs. Hazleton Food Science Newsletter, 2 (38):1.

Nutrition Policy Board, USDHHS; (1988) The Surgeon General's Report on Nutrition and Health. U.S. Department of Health and Human Services, Public Health Service, DHHS (PHS) Publication No. 88-50210. U.S. Government Printing Office, Washington, DC 20402.

Committee on Diet and Health, Food and Nutrition Board; (1989) Diet and Health: Implications for Reducing Chronic Disease Risk. Commission on Life Sciences, National Research Council, National Academy of Sciences, National Academy Press, Washington, D.C.

Industry Response to Labeling Initiatives:

# Impact on Industry Activities Regarding Food Analysis and Nutrient Databases

**Frances H. Seligson, Ph.D., R.D.**
**Hershey Foods Corporations**
**Hershey, Pennsylvania**

## INTRODUCTION

This presentation: (1) covers current practices at Hershey Foods for providing nutrition information to consumers and professionals; (2) addresses the impact of NLEA (The Nutrition Labeling and Education Act of 1990) and FDA (Food and Drug Administration) proposed regulations on these practices; and (3) highlights some concerns related to these labeling initiatives.

## CURRENT PRACTICES

Hershey Foods has provided nutrition information on a voluntary basis for the majority of its products since 1973. For products manufactured by HCUSA (Hershey Chocolate USA), nutrition information is provided on the label for 75% of our products on a sales volume basis (Table 1). It is mostly our confectionery products which carry nutrition label information and this information usually includes a voluntary declaration of total sugar. Nutrition information is available upon request for another 21% of our products, mostly our grocery products (i.e., baking chocolate, baking chips, and syrup), GOLDEN chocolate confections, and nonchocolate confections. The food label for many of these products contains our consumer 800 number; and the nutrition information which is available is based on a combination of analytical and calculated data. For the remaining 4% of HCUSA products, mostly low volume and/or seasonal items, nutrition information is not readily available but estimates can be obtained upon request.

HPG (Hershey Pasta Group) manufactures a wide variety of regional, branded pasta products, which account for 90% of its product sales (Table 2). All of

TABLE 1
**Nutrition Information for Hershey Chocolate USA Products**

| | *% of Sales* |
|---|---|
| Provided on food label[1] | 75 |
| Available upon request[2] | 21 |
| None available[3] | 4 |

[1]Confectionery; analytical data
[2]Grocery; analytical and calculated data
[3]Low volume/seasonal items

TABLE 2
**Nutrition Information for Hershey Pasta Group Products**

| | *% of Sales* |
|---|---|
| Provided on food label[1] | 90 |
| None available[2] | 10 |

[1]Branded, enriched pasta
[2]Rice, dried beans, sauces, bread crumbs

the enriched products must and do carry nutrition information on the food label. HPG also sells instant mashed potatoes, dried rice and various legumes, sauces, and bread crumbs under regional brand names; and nutrition information typically is not available for these products.

Hershey Foods also provides nutrient information for confectionery and grocery products in brochure form upon request (Table 3). Our *Nutrition Information for Consumers* brochure lists nutrient data per serving in the format in which it is provided on the food label for most of our confectionery and grocery products. Cholesterol and total sugars are also provided. The information in this brochure is compiled from several sources, viz., product analysis and calculations using formulations and nutrient databases.

TABLE 3
**Nutrient Information Brochures and Data Sheets**

- Nutrition Information for Consumers[1]
- Nutrition Information for Health Professionals[2]
- Cholesterol and Fatty Acid Information[3]

[1]per serving; nutrition label format; cholesterol and sugar
[2]per 100 grams; analytical average; cholesterol, sugar, added minerals
[3]per serving and per 100 grams

Another brochure, *Nutrition Information for Health Professionals*, provides nutrition information per 100 grams for only those confectionery and grocery products with analytical data. In addition to those nutrients required on the food label under current regulations, information is presented for cholesterol, copper, magnesium, manganese, phosphorus, potassium, and zinc. It is this information that we provide to individuals who are building nutrient databases.

Cholesterol and fatty acid information is also available upon request for about 23 products manufactured by HCUSA. The information is provided both per serving, rounded according to nutrition labeling rules, and per 100 grams as the analytical average.

In compiling nutrient data for HCUSA confectionery and grocery products, we ideally like to use analytical results from seven production lots (each lot is represented by a composite of 12 samples) for the information which appears on the label. This number of analyses gives us a high level of statistical confidence that our label information would be in compliance if challenged. For those nutrients which we provide voluntarily in brochures, we like to use the analytical results from at least three production lots. It costs about $3,870 per product to obtain this amount of analytical information (Table 4).

TABLE 4
**Nutrient Information Analytical Costs: Current Practices**

| | Cost per Production Lot Sample | Number Production Lots | Total Cost |
|---|---|---|---|
| Current Label Requirements | $435 | 7 | $3,045 |
| Cholesterol & Fatty Acids | 195 | 3 | 585 |
| Additional Minerals | 80 | 3 | 240 |
| | | | $3,870 |

## IMPACT OF NLEA

The NLEA and FDA proposed regulations will definitely impact our current practices (Table 5). First and foremost, we have to develop nutrition label information for our entire product line, and we are currently in the process of taking inventory of our nutrient label database. We are trying to determine

TABLE 5
**NLEA Impact on Current Practices**

- All products will be labeled
- Fewer production lots will be analyzed
- Fewer optional nutrients will be analyzed
- Increased workload and staffing
- Upgrade computer system

what is needed in terms of product analyses, nutrient analyses, and costs to meet the demands of mandatory nutrition labeling. The cost incurred with obtaining analytical data for the basic nutrition label information required under NLEA and FDA proposed regulations will increase by about $160 per product sample (Table 6), so we may decide to rely on fewer production lots to construct the basic nutrition label. We may also decide not to analyze for nutrients that are not likely to occur in a significant amount. Similarly, we will probably analyze fewer products for optional nutrients, if we continue with this practice at all.

TABLE 6
**Nutrient Information Analytical Costs for Required Nutrients**

| | *Cost per Production Lot Sample* |
|---|---|
| Current label requirements | $435 |
| Proposed label requirements[1] | 595 |
| Difference | $160 |

[1]NLEA plus vitamins A & C, calcium, and iron

Another major area of impact has been on our workload. We have increased our corporate nutrition labeling support staff from 0.6 FTE (full time equivalent) to 2 FTE's (this does not include the effort that has been placed against tracking and addressing the various labeling initiatives). These FTE's are handling ongoing requests for nutrition information, assessing what needs to be done for mandatory labeling, and upgrading our computerized system to handle an expanded nutrient database. The goal is to electronically store and share nutrient information with HCUSA and HPG Quality Assurance Departments.

## CONCERNS

We have several concerns in responding to the labeling initiatives, the biggest one of which is time. There are only six months between the time final regulations are published and when they will be enforced. We are reluctant to analyze products until we are certain of the details of what will be required, e.g., how to analyze for fiber, sugar, and complex carbohydrates.

Related to time, we are concerned about the ability of contract analytical laboratories to accommodate the increased demand for services. We also have concerns about the reliability of the data which will be used to construct the label. First, we will probably rely on fewer production lots, and second, some of our product matrices (especially chocolate and coconut) are analytical challenges. Lastly, we are concerned there will be less time available for responding to requests for nutrition information from the people who build nutrient databases. But that is a topic for another session!

TABLE 7

**Concerns About NLEA**

- Time for enforcement
- Contract lab capability
- Reliability of data
- Increased workload

Industry Response to Labeling Initiatives: Impact on Industry Activities

# Nutrient Databases for the 90's

**Katy Raneri, M.S., R.D.**
**Kraft General Foods**
**Tarrytown, New York**

## I. Overview/History

The GF USA Nutrient Database has been in existence since the early 1970s, developed primarily for use in nutrition labeling. It has undergone numerous transformations over the years, earning the name NUTRIFILE in 1983.

NUTRIFILE's major functions include the support of:

* Nutrition label claims, including substantiating nutrition and health claims
* Product development, by enabling product developers to design products around specific nutrient criteria
* Recipe and special diet system development

NUTRIFILE carries out these tasks by:

1. Calculating the nutrient profiles of product formlas, recipes and raw materials.
2. Storing and summarizing analytical data.
3. Performing adjusted analytical analyses - calculation of the theoretical impact of a change in product formulation using previously obtained analytical data.
4. Generation of a variety of reports - compositional, comparison, audit, etc.
5. Analyzing dietary intakes - calculates and stores the nutrient profiles of a diet record using age/sex appropriate RDAs.

NUTRIFILE currently contains nutrient profiles for over 15,000 food items, including GF products, ingredients and generic foods. 35 nutrients are stored for each food item, with the capacity to store up to 95 nutrients.

## II. Meeting Greater Business Needs

The diversity, complexity and everchanging nature of our business, coupled with rapidly expanding computer technology, made it necessary to re-evaluate and ultimately restructure our system. We are currently redesigning and enhancing our database. The new system is pc-based and operates on a Local Area Network. The objectives of our enhanced database include:

* Increased functionality to address current and anticipated business needs and requirements.
* Improved productivity by significantly reducing annual operating costs.
* Flexibility for growth.

In addition, all these objectives must be met while continuing to meet on-going business needs.

## III. Challenges of a Sophisticated Nutrient Database

We believe we have a state-of-the-art nutrient database and the challenges of a sophisticated system are many. Along with all the technical and business requirements necessary to achieve a state-of-the-art system, there is the external environmental challenge as presented in the 1990 Food Labeling Reform. Some of the challenges which we, as well as many of you, are currently facing include:

* **Nutrients/Data**

  - what nutrients will become mandatory and how will they be defined?

  - availability of data - there isn't sufficient fiber or complex carbohydrate data

  - nutrient units (i.e. RE vs IU for vitamin A)

  - format - certainly any format changes will impact how the nutrients are presented

* **Advances in Food Technology**

Although having analytical data for a food product is more appropriate in certain situations (unique new product development, new or novel ingredients/processes, extensive/unfamiliar processing, multiple components) we can certainly learn a great deal about mirroring these processing changes by also performing calculations and comparing the results. Gone are the days when a product portfolio consists of only "simple" ingredients of fairly constant composition, manufactured by mixing in the absence of thermal processing.

* **Advances in Computer Technology**

  - The exponential growth in the power of the pc

  - Graphical interfaces give a quite different look and feel to our system

  - Advances in computer technology are occuring every day and we have to realize that quite likely our final system will be somewhat outdated by the time we complete the project.

* **Expertise of User**

  - New and novel processing
  - Advances in computer technology
  - Advances in nutrition
  - Regulatory environment

All of the above are a challenge individually. When you consider all of them together, the challenge increases considerably. The importance of user training, knowledge and expertise cannot be overstated. A database is only as good as the user.

Industry Response to Labeling Initiatives:
Impact on Industry Activities

# Nutrient Databases for the 90's

Katy Raneri, M.S., R.D.
Kraft General Foods
Tarrytown, New York

## I. Overview/History

The GF USA Nutrient Database has been in existence since the early 1970s, developed primarily for use in nutrition labeling. It has undergone numerous transformations over the years, earning the name NUTRIFILE in 1983.

NUTRIFILE's major functions include the support of:

* Nutrition label claims, including substantiating nutrition and health claims
* Product development, by enabling product developers to design products around specific nutrient criteria
* Recipe and special diet system development

NUTRIFILE carries out these tasks by:

1. Calculating the nutrient profiles of product formlas, recipes and raw materials.
2. Storing and summarizing analytical data.
3. Performing adjusted analytical analyses - calculation of the theoretical impact of a change in product formulation using previously obtained analytical data.
4. Generation of a variety of reports - compositional, comparison, audit, etc.
5. Analyzing dietary intakes - calculates and stores the nutrient profiles of a diet record using age/sex appropriate RDAs.

NUTRIFILE currently contains nutrient profiles for over 15,000 food items, including GF products, ingredients and generic foods. 35 nutrients are stored for each food item, with the capacity to store up to 95 nutrients.

## II. Meeting Greater Business Needs

The diversity, complexity and everchanging nature of our business, coupled with rapidly expanding computer technology, made it necessary to re-evaluate and ultimately restructure our system. We are currently redesigning and enhancing our database. The new system is pc-based and operates on a Local Area Network. The objectives of our enhanced database include:

* Increased functionality to address current and anticipated business needs and requirements.
* Improved productivity by significantly reducing annual operating costs.
* Flexibility for growth.

In addition, all these objectives must be met while continuing to meet on-going business needs.

## III. Challenges of a Sophisticated Nutrient Database

We believe we have a state-of-the-art nutrient database and the challenges of a sophisticated system are many. Along with all the technical and business requirements necessary to achieve a state-of-the-art system, there is the external environmental challenge as presented in the 1990 Food Labeling Reform. Some of the challenges which we, as well as many of you, are currently facing include:

* **Nutrients/Data**

  - what nutrients will become mandatory and how will they be defined?

  - availability of data - there isn't sufficient fiber or complex carbohydrate data

  - nutrient units (i.e. RE vs IU for vitamin A)

  - format - certainly any format changes will impact how the nutrients are presented

* **Advances in Food Technology**

Although having analytical data for a food product is more appropriate in certain situations (unique new product development, new or novel ingredients/processes, extensive/unfamiliar processing, multiple components) we can certainly learn a great deal about mirroring these processing changes by also performing calculations and comparing the results. Gone are the days when a product portfolio consists of only "simple" ingredients of fairly constant composition, manufactured by mixing in the absence of thermal processing.

* **Advances in Computer Technology**

  - The exponential growth in the power of the pc

  - Graphical interfaces give a quite different look and feel to our system

  - Advances in computer technology are occuring every day and we have to realize that quite likely our final system will be somewhat outdated by the time we complete the project.

* **Expertise of User**

  - New and novel processing
  - Advances in computer technology
  - Advances in nutrition
  - Regulatory environment

All of the above are a challenge individually. When you consider all of them together, the challenge increases considerably. The importance of user training, knowledge and expertise cannot be overstated. A database is only as good as the user.

# Nutrient Data Resources Available from HNIS

**David B. Haytowitz**
**USDA / HNIS**
**Hyattsville, Maryland**

This paper will describe a number of the various nutrient data bases developed by HNIS, what's in them, and how to obtain them.

All of the information in this paper is also available on the Nutrient Data Bank Bulletin Board. In addition, some of the files described herein can also be downloaded from the Bulletin Board. The Bulletin Board was first announced at this conference in 1989 and was demonstrated at the conference last year in Blacksburg, Virginia. Since its inception, the Bulletin board has received almost 2,000 calls from 44 states and Canada.

In order to access the Bulletin Board you need an IBM compatible PC, a modem, and a communications program. The telephone number is (301) 436-5078, and it is available 24 hours a day, 7 days a week, except for brief periods for maintenance. Therefore you can call when the phone rates are lower. We plan to update the bulletins, as appropriate, during the first week of each month. Of course, new information will be posted immediately. While the Bulletin Board can be accessed from mainframes and other computers, you won't be able to uncompress any files you may download, because they are compressed using a MS-DOS program. However, you will be able to read all of the bulletins.

There are also a number of small files available for downloading. These files will be described below. In addition to the data files, the Dietary Analysis Program developed by HNIS and the Extension Service can also be downloaded from the Bulletin Board. The data contained in our provisional tables on vitamins D and K are also available. As new provisional tables are released, data sets will be added to the Bulletin Board. You can also leave messages for the sysop. Since we have not implemented the full message capabilities of the board, we will contact you by phone, so please leave a number where you can be reached.

## USDA NUTRIENT DATA BASE FOR STANDARD REFERENCE

The USDA Nutrient Data Base for Standard Reference (or Standard Reference) is the machine readable version of Agriculture Handbook No. 8. The current version is Release 9 which was issued last summer (1990) when the revised data on beef in the 1990 revision of AH-8-13 was published. It contains data from sections 1-17, 20, 21 and the 1989 supplement (Table 1). Data from the 1963 edition of AH-8 is retained for unrevised sections. Earlier data are removed as new data from the revised sections are added to the data base. It is expected that most of these data will be replaced in the near future. There is also an abbreviated version of this data base which contains fewer nutrients but all of the food items.

Release 10, which will add data from the second (1990) and third (1991) supplements as well as data from AH-8-19 on Snacks and Sweets and possibly data from AH-8-18 on Baked Products, will be available later this year or early 1992. The third (1991) supplement will contain new pages for a major revision of AH-8-10 on Pork Products. Standard Reference uses the NDB numbers, which are printed at the bottom of each page of the handbook to identify the food items. The first two digits (01-21) denote the food group. Food groups 22 and 23 which had previously been identified as "Mixed Dishes" and "Miscellaneous" have been dropped. Items slated for inclusion for these sections will be included in the remaining sections and in supplement pages for previously issued sections. The last three digits denote the food item within the group.

The Standard Reference Data Base is available on both diskettes and magnetic tape. Diskettes are formatted for IBM-compatible machines and are available in 3-1/2" and 5-1/4" formats, both low and high density.

Table 1 - **Agriculture Handbook No. 8 Series**

| Series No. | Food group | Year Issued | No. of Items |
|---|---|---|---|
| 8-1 | Dairy and Egg Products | 1976 | 144 |
| 8-2 | Spices and Herbs | 1977 | 43 |
| 8-3 | Baby Foods | 1978 | 217 |
| 8-4 | Fats and Oils | 1979 | 128 |
| 8-5 | Poultry Products | 1979 | 304 |
| 8-6 | Soups, Sauces, and Gravies | 1980 | 214 |
| 8-7 | Sausages and Luncheon Meats | 1980 | 80 |
| 8-8 | Breakfast Cereals | 1982 | 142 |
| 8-9 | Fruits and Fruit Juices | 1982 | 263 |
| 8-10 | Pork Products | 1983 | 186 |
| 8-11 | Vegetables and Vegetable Products | 1984 | 470 |
| 8-12 | Nut and Seed Products | 1984 | 117 |
| 8-13 | Beef Products | 1990 | 367 |
| 8-14 | Beverages | 1986 | 153 |
| 8-15 | Finfish and Shellfish Products | 1987 | 174 |
| 8-16 | Legumes and Legume Products | 1986 | 133 |
| 8-17 | Lamb, Veal, and Game Products | 1989 | 221 |
| 8-18 | Baked Products | In preparation | |
| 8-19 | Snacks and Sweets | 1991 | 301 |
| 8-20 | Cereal Grains and Pasta | 1988 | 123 |
| 8-21 | Fast Foods | 1988 | 166 |
| 1989 | Supplement | 1990 | 96* |
| 1990 | Supplement | 1990 | 114* |
| 1991 | Supplement | In preparation | |

* New and revised items

Table 2 - **Nutrient in the USDA Nutrient Data base for Standard Reference.**

| Proximates | Minerals | Vitamins[3] |
|---|---|---|
| Water | Calcium | Ascorbic acid |
| Protein | Iron | Thiamin |
| Fat | Magnesium | Riboflavin |
| Carbohydrate[1] | Phosphorus | Niacin |
| Crude fiber[2] | Potassium | Pantothenic acid |
| Energy | Sodium | Vitamin $B_6$ |
| | Zinc | Folacin |
| | Copper[4] | Vitamin $B_{12}$ |
| | Manganese[5] | Vitamin A (IU and RE) |

| Lipids | Amino acids |
|---|---|
| Fatty acids | Tryptophan |
| Total saturated | Threonine |
| Total monounsaturated | Isoleucine |
| Total polyunsaturated | Leucine |
| Individual fatty acids | Lysine |
| Cholesterol | Methionine |
| Plant sterols | Cystine |
| | Phenylalanine |
| | Tyrosine |
| | Valine |
| | Arginine |
| | Histidine |
| | Alanine |
| | Aspartic acid |
| | Glutamic acid |
| | Glycine |
| | Proline |
| | Serine |

[1] By difference.
[2] If available, total dietary fiber is given.
[3] If available, alpha-tocopherol is given.
[4] Not available for handbook sections 8-4.
[5] Not available for handbook sections 8-3 and 8-4.

For those of you who have purchased a previous release and do not want to replace all of the data when a new release is issued, an update is available. These updates contain all of the information needed to update the previous release. For example, the update to Release 8 contains the new data on Beef Products published in AH-8-13 (1990), as well as a few corrections. Once you have updated your copy of Release 8 with these data, you will have a data set identical to Release 9. The update files can be purchased from NTIS on magnetic tape and all diskette formats described above. The update files are also available on the Bulletin Board.

Release 9 of the Standard Reference Data Base contains data on 5,145 food items. For the revised section of AH-8 it contains all of the nutrients listed in Table 2. Any proximate, mineral, or vitamin values missing on the AH-8 printed pages were imputed for the tape. Imputed values are calculated from either another form of the same food or from a similar food. Also included on the data file is a brief (20 characters) description of the food item. A more complete (72 characters) description is contained in a separate file, which also contains the column headings and weights for columns E, F, and G from the printed AH-8 pages. The microcomputer version contains factors for calculating the values presented in columns E, F, and G in order to save space, while the magnetic tape version contains the calculated values.

## DATA SET 72-1

Data set 72-1 contains the nutrient data from Home and Garden Bulletin No. 72. Latest version is Release 3.2 which contains the updated values on eggs, which were released in our 1989 supplement, as well as a few corrections. These data are expressed in terms of common household units. The data set is available on diskettes (same formats and densities as Standard Reference) and magnetic tape. This data set can also be downloaded from the Bulletin Board.

The data set contains information on 961 food items, arranged by food groups. The printed publication contains an index. Each item is assigned an unique 4-digit number. The data set also contains a description of each food and the household weight. The nutrients in this data set are listed in Table 3.

Table 3 - **Nutrients in Data Set 72-1.**

| Proximates | Minerals | Vitamins |
|---|---|---|
| Water | Calcium | Ascorbic acid |
| Protein | Iron | Thiamin |
| Fat | Phosphorus | Riboflavin |
| Carbohydrate[1] | Potassium | Niacin |
| Energy | Sodium | Vitamin A (IU and RE) |

| Lipids |
|---|
| Fatty acids |
| Total saturated |
| Total monounsaturated |
| Total polyunsaturated |
| Cholesterol |

[1] By difference.

## USDA NUTRIENT DATA BASE FOR FOOD INTAKE SURVEYS

The USDA Nutrient Data Base for Food Intake Surveys (or Survey Nutrient Data Base) was first developed for the 1977-78 Nationwide Food Consumption Survey and contained data on 15 nutrients.

This was Release 1 of the data base.Release 2 was developed for the Continuing Survey of Food Intakes by Individuals (CSFII) and was used for the first set of data collected in the 1985 survey (Wave 1 core monitoring group). Release 2.1 covers the complete 1985 survey and contains approximately 500 additional foods for approximately 4,500 items.

Release 3 was developed for the 1986 CSFII and contains approximately 5,300 items. Release 4 was developed for the 1987-88 Nationwide Food Consumption Survey and contains approximately 6,300 items.

Releases 2.0, 2.1, 3 and 4 all share the same format and contain the nutrients listed in Table 4. Each item is identified by a 7-digit food code. Contained in each food record is a 51-character description of the food item. The record also contains a "fat in cooking code." This code is used to access the nutrient records calculated using fats or oils other than the one designated in the recipe for a particular item. For example, if butter were the designated fat in the recipe file, an alternate nutrient profile would be calculated for the same food cooked in margarine and several other cooking fats and oils. A "salt in cooking code" is used in those recipes where the food preparer has the choice of adding or not adding salt to the recipe. In these cases two records are created--one with salt added and the other with no salt added.

Data from Release 4 is available on the Bulletin Board. However, to save space only the default fat values are included. Furthermore, the two salt records are combined in one record with two sodium fields. If salt is added to a recipe, one field contains the sodium value for salt added and the other field contains the sodium value if no salt was added. If the recipe does not contain salt or if the food preparer does not have the option of omitting salt from the recipe, both sodium fields contain the same value. Also included is the survey codebook, which contains full descriptions and household weights for each item in the Survey Nutrient Data Base.

## DATA SETS USED TO CREATE THE USDA NUTRIENT DATA BASE FOR FOOD INTAKE SURVEYS.

There are three data sets that contain the data used to create the USDA Nutrient Data Base for Food Intake Surveys. They are (1) Primary Nutrient Data Set for USDA Nationwide Food Consumption Surveys, (2) USDA Table of Nutrient Retention Factors, and (3) Recipe File for the USDA Nutrient Data Base for Individual Food Intake Surveys.

The Primary Nutrient Data Set for USDA Nationwide Food Consumption Surveys (PDS) contains the nutrient data used to calculate the nutrient values in the Survey Nutrient Data Base. It contains the same 30 nutrients as the Survey Nutrient Data Base and is based on the Standard References releases as follows:

| Survey | PDS | Standard Reference |
|---|---|---|
| CSFII 85 | 1 | 5 |
| CSFII 86 | 2 | 5 |
| NFCS 87-88 | 3 | 7 |

Data on the nutrient content of eggs and beef in Release 3 were adjusted to reflect changes in the food supply. If a nutrient value was missing from the Standard Reference (either because that nutrient is not in Standard Reference or a value for a nutrient that is in Standard Reference is missing) it was added to the PDS. In addition, new foods not in the matching release of NDB-SR were also added as needed. When a new release of Standard Reference contains a food previously added to the PDS, that item is replaced with the new values from the latest release of the Standard Reference when the PDS was updated for a new survey. Future releases of this database will contain additional foods as well as updates from new releases of the Standard Reference. For example, we are currently working to update the PDS to include data from Releases 8 and 9 of Standard Reference. The PDS contains the same nutrients as the Survey Nutrient Data Base, which are listed in Table 4. The PDS also contains a brief (20 characters) description. A longer (72 characters) description appears in a separate file.

The USDA Table of Nutrient Retention Factors contains the retention factors used to calculate the nutrient values in the Survey Nutrient Data Base. The factors are based primarily on the HNIS 1984 "Provisional Table on Percent Retention of Nutrients in Food Preparation." Additional factors and food categories have also been added to the table. Releases of this table are as follows:

| Survey | Retention Factors |
|---|---|
| CSFII 85 and 86 | Release 1 |
| NFCS 87-88 | Release 2 |

There are retention factors for 16 minerals and vitamins. If a factor is not present in this table, the program used to calculate the nutrient values in the Survey Nutrient Data Base assumes 100% retention. The table contains a 4-digit computer code that is used to access the retention data as well as a brief description. The retention factors are reviewed periodically to insure that the most up-to-date values, based on the latest research, are used.

The last of the three files used to create the Survey Nutrient Data Base is the Recipe File for the USDA Nutrient Data Base for Individual Food Intake Surveys. This file, along with a computer program, ties the values in the PDS and retention factors file together to create the Survey Nutrient Data Base. The recipe file contains the component ingredients and their proportions, plus a retention code if applicable. There is a recipe record for every item on the Survey Nutrient Data Base, although approximately one-half of these recipes are single ingredient recipes--i.e, they only reference a single PDS record. The file contains a header record for each recipe. In addition to the name of the recipe, the header record also contains yield, including if applicable moisture loss or gain, and fat loss or gain. The ingredient records contains the PDS identification number, a description (used as a check on the data entry), a retention factor code, the measure (optional; entered for documentation), and the weight or proportion of that ingredient. There are also codes to indicate if alternate fat and salt records are to be calculated for the Survey Nutrient Data Base. The PDS code can be replaced with a code designating another recipe. For example to calculate the nutrient values for spaghetti with meat sauce, the 7-digit food code for meat sauce, which was calculated previously as a separate item in the Survey Nutrient Data Base, is combined with the code for cooked spaghetti.

A documentation file accompanies these three files. It contains (1) documentation for the three files, (2) a description of the recipe calculation procedure, and (3) a coding manual for the PDS. Ordering instructions for data sets and other data sets prepared by the Human Nutrition Information Service are available from our office. The same information is also available on the Nutrient Data Bank Bulletin Board.

Table 4 - **Nutrients in the USDA Nutrient Data Base for Individual Food Intake Surveys and Primary Nutrient Data Set for USDA Nationwide food Consumption Surveys.**

| Proximates | Minerals | Vitamins |
|---|---|---|
| Water | Calcium | Ascorbic acid |
| Protein | Iron | Thiamin |
| Fat | Magnesium | Riboflavin |
| Carbohydrate[1] | Phosphorus | Niacin |
| Total Dietary Fiber | Potassium | Vitamin $B_6$ |
| Energy | Sodium | Folacin |
| Alcohol | Zinc | Vitamin $B_{12}$ |
| | Copper | Vitamin A (IU and RE) |
| | | Carotene |
| | | Vitamin E |

**Lipids**

Fatty acids
- Total saturated
- Total monounsaturated
- Total polyunsaturated

Cholesterol

[1] By difference.

# New Users: Nutrient Database Features

**R. Sue McPherson, Ph.D.**
**University of Texas**
**Houston, Texas**

A nutrient database is a collection of interrelated data about foods and constituents of foods stored together without unnecessary redundancy to serve one or more applications in an optimal fashion. These data are independent of software programs that access the data. A nutrient database may also have data included that are related to the foods which are used for identification and documentation.

The major components of a nutrient database system include: 1) a nutrient database; 2) access and calculation software; and 3) reporting software. Selecting a nutrient database system requires careful evaluation of each of these components in relationship to the specific goals and objectives of the user's work environment. Some questions to investigate when evaluating nutrient database systems are listed below:

1. What is the primary data source?
2. What nutrients are included?
3. How many foods are included?
4. Are there mixed dishes, recipes or ethnic foods?
5. How are missing values handled?
6. How are additions to the primary database done?
7. How is the database maintained?
8. What documentation is available?
9. Is there user support?
10. Are custom services available?
11. What hardware does the program require?
12. What are the data storage capabilities?
13. What is the response time?
14. What features does the software have?
15. How is data entry completed?
16. What analysis options are there?
17. Are there comparison data?
18. Can custom features be added?
19. How many users does the software support?
20. What printers are supported?
21. What are the costs?
22. Does the system meet the user needs and objectives?

This list of questions may be used as an aid in exploring available nutrient database systems. It is important to always separate the evaluation of the quality and quantity of the nutrient database from the evaluation of the features of the software programs which access these data. Although it is easy to be distracted in the review of software applications which streamline the access and reporting tasks, remember that the credibility of reports are dependent on the quality of the nutrient database.

In summary, the assessment of dietary intake using any methodology is dependent on having a quality nutrient database. The selection of a nutrient database requires the knowledge of the goals and objectives of the project. As well, selection of the software to utilize a nutrient database requires careful evaluation in order to achieve project goals. Finally, the development of a nutrient database system is a commitment to the future, and requires long-range planning and maintenance.

# Update on AOAC Analytical Methods for Nutrients

**William Horwitz**
**Food and Drug Adminstrator**
**Washington, District of Columbia**

During the past decade, the development of "new" methods of analysis for nutrients, or even the application of what are now standard methods for other substances, has been limited. Most of the activity has been in applying previously approved methods for nutrients in foods to infant formulas. From the point of view of database production, perhaps such stability is a desirable attribute. The purpose of a database is to solidify, coordinate, and preserve knowledge; one of the worst things that can happen to a compiler is a slow, unrecognized drift of what was thought to be a stable reference point.

Over the last half century, chemical, microbiological, and bioassay procedures have been developed for nutrients, and many of them have been validated by the application of method-performance protocols of the Association of Official Analytical Chemists (AOAC). The AOAC has been testing and adopting methods of analysis required to enforce laws and regulations for over a century. Its methods are accepted professionally and legally on an international basis. AOAC relies upon its sponsoring agencies, primarily to US Food and Drug Administration, the US Department of Agriculture, and the corresponding organizations in Canada, other countries, and the states for the actual performance of the laboratory studies required to gain approval of methods as "official."

Many of the currently approved methods for nutrients were validated 20 to 40 years ago and have not been updated to take advantage of modern separation and measurement technology. For example, validated automated methods for nutrients exist only for niacin, riboflavin, and vitamin C. Although innumerable methods exist in the literature proposing "new and improved" high-performance liquid chromatographic methods for the fat-soluble vitamins in foods -- vitamin A, vitamin D, and vitamin E – such a method has been adopted only for vitamin D in milk, milk powder, and mixed feeds. The structural components of food, primarily protein and fiber-related components, appear to interfere with the release of the minor and trace nutrients to the measurement operations.

Although a vast literature exists on methodology for nutrients, very few methods have gone to an interlaboratory methods-performance trial. Many methods that behave admirably in the hands of the proud developers fail miserably when subjected to the abuses inflicted by other laboratories. Part of the problem of development of new methods for nutrients lies with the lack of stable reference standards to ensure a common reference point for all laboratories. Furthermore, knowledge of the variability of the nutrient concentrations across laboratories is important to indicate the relative standing of analytical variability compared to the variability introduced from other sources such as physical sampling of raw commodities (processing tends to smooth out commodity variability), uncontrolled serving sizes, and total food intake of individuals.

Table 1 provides the official AOAC method number for nutrients in the latest edition (1990) of *Official Methods of Analysis of the AOAC,* a reference to the last interlaboratory method-performance study published in the *Journal of the Association of Official Analytical Chemists,* and a brief personal comment on the quality of the results. The high variability of published results for nutrients that exist in the literature strongly suggests that little attention has been paid in the past to quality control of the analytical operations. It is possible that a considerable fraction of the variability ascribed to varieties, geography, and seasonal factors of raw commodities is merely a manifestation of lack of quality control of the nutrient analyses. In the single commodity area of infant formulas, which had the benefit of legislation mandating adherence to nutrient standards, a decade of investigational work is expected to result, at the AOAC meeting of August 1991, in a complete set of methods for measuring all of the nutrients required to be stated on the label of this commodity.

Table 1. Reference to Analytical Methods in the Vitamins and Other Nutrients Chapter in the 15th Edition (1990) of **Official Methods of Analysis of the Association of Official Analytical Chemists**

| Analyte / Matrix | Method Type | Method Number* | Latest JAOAC Reference |
|---|---|---|---|
| | ***Chemical Methods*** | | |
| **Vitamin A** | | | |
| Foods | Colorimetric | 974.29 | 63,0468 (1980) |
| | The official method is still the antimony trichloride spectrophotometric procedure. Carotenes can be determined simultaneously in high fat products. Note there is no general HPLC method for Vitamin A. | | |
| Carotenes in Fresh Plants | Spectrophotom | 941.15 | 53,0186 (1970) |
| | Although monohydroxy and dihydroxy pigments are reported, there is no provision for these analytes in the referenced method, 941.15. | | |
| **Thiamine** | | | |
| Foods | Fluorometric | 942..23 | 64,1336 (1981) |
| Grains | Fluorometirc | 953.17 | 38,0722 (1955) |
| Bread | Fluorometric | 957.17 | 43,0047 (1960) |
| | All three methods are based upon the classical thiochrome fluorometric measurement. A number of simplifications are available for enriched foods, in which the thiamine is present in free form and does not have to be released from irs compounds. | | |
| **Riboflavin** | | | |
| Foods | Fluorometric | 970.65 | 53,0542 (1970) |
| | Latest version was also applied to concentrates and feed supplements. | | |
| | Automated Flourometric | 981.15 | 62,1041 (1979) |
| | Extensive study but results from one laboratory are somewhat out of line. | | |
| **Niacin and Niacinamide** | | | |
| Foods | Colorimetric | 961.14 | 45,0449 (1962) |
| | Excellent parameters but only for bran, bread, cereal, and flour. | | 44,0431 (1961) |
| Cereals | Automated Colorimetric | 975.41 | 58,0799 (1975) |
| | Automated method much superior to manual method; microbiological method highly variable as applied to cereals and baked pet food. | | |
| Foods | Automated Colorimetric | 981.16 | 62,1027 (1979) |
| | Manual method erratic; automated method much superior. Even among-laboratories precision better for automated method than for manual or microbiological method for a variety of foods. | | |
| **Vitamin C (Ascorbic Acid)** | | | |
| Juices | Titrimetric | 967.21 | 50,0798 (1967) |
| Foods | Automated Fluorometric | 984.26 | 66,1371 (1983) |
| | All assays, except those by the unofficial diphenylhydrazine method, are excellent. | | |
| **Vitamin D** | | | |
| Milk and Powder | HPLC | 981.17 | 65,1228 (1982) |
| | Variability high but acceptable. | | |
| Mixed Feeds | HPLC | 982.29 | 66,0751 (1983) |
| | Variability high but acceptable. Thios aasy requires conscientious quality control. Ont the practice test sample, the range of 18 laboratoriaes was a factor of 4; on real test samples, the range was 5-10. One laboratory reported 146 for a 3 ppm concentration. | | |

* The Method Number is constructed from the last three digits of the year of adoption and the consecutive item number separated by a decimal point.

| Analyte Matrix | Method Type | Method Number* | Latest JAOAC Reference |
|---|---|---|---|
| **Vitamin E** | | | |
| Alpha-tocopherol (AT) | Colorimetric | 971.30 | 54,00001 (1971) |
| | Obsolete TLC method; apparently problems exist in nomenclature and units. | | |
| **Calcium Pantothenate** | | | |
| | Spectrophotometric | *945.73* | 52,0449 (1969) |
| | The only chemical method is for pharmaceutical preparations. | | |
| **Sodium** | | | |
| Foods for SDU | Ion Sel. Electrometric | 976.25 | 59,1131 (1976) |
| | Ion selective electrode method is very simple, but lower limit is about 10 ppm Na. Precautions necessary to avoid ordinary laboratory contamination. | | |
| **Fat** | | | |
| Foods | $CHCl_3$-MeOH Extn | 983.23 | 66.0927 (1983) |
| | Fat values are method-specific; individual commodities have classsical methods, but composite products use MeOH-$CHCL_3$ and must be thoroughtly homogenized. Relative standard deviations as haigh as 20% can be expected at 1-5% fat levels. | | |
| **Cholesterol** | | | |
| Multicomponent Foods | GLC | 976.26 | 59,0046 (1976) |
| | Precision parameters relatively poor apparently from both homogenization and fat extraction. | | |
| Noodles | Fluorometric | 969.14 | 51,1220 (1968) |
| | Very good parameters but limited to noodles. | | |
| **Dietary Fiber** | | | |
| Foods | Enzymatic-Grav. | 985.29 | 69,0259 (1986) |
| | Problems with this analyte are very well known.. Values are method-specific, and reproducibility at 1-5% levels and below is very poor. | | |
| **Protein, carbohydrates, sugars, etc.** | | | |
| | Numerous methods available for individual commodities. See overall evaluation in 73,0661 (1990) and for dairy products in 72,0784 (1989). | | |
| | ***Microbiological Methods*** | ***Titrimetric*** | ***960.46E*** |
| | | ***Tubidimetric*** | ***960.46F*** |
| **Cobalamin** | | | |
| $B_{12}$ Activity | Microbiological | 952.20 | 45,0529 (1959) |
| | Studies performed on relatively high activity products with excellent results. | | |
| **Folic Acid** | Microbiological | 944.12 | 42,0529 (1959) |
| | Studies performed on relatively high activity products with excellent results. | | |
| **Niacin and Niacinamide** | Microbiological | 944.13 | 62,1027 (1959) |
| | Ten commodities as part of the automated method study provided acceptable results. | | |
| **Pantothenic Acid** | Microbiological | 945.74 | 42,0853 (1957) |
| | Excellent results but only for three products and by five laboratories. | | |
| **Riboflavin** | Microbiological | 940.33 | 32,0461 (1949) |
| | Acceptable results but only a few test samples. | | |
| **Amino Acids** | Microbiological | 960.47 | 43,0034 (1960) |
| | Excellent results but the studies were conducted with solutions. | | |
| **Vitamin B6** | | | |
| Pyridoxine, -al, --amine | Microbiological | 961.15 | 53,0546 (1970) |
| | Results are acceptable for most foods, but liver is exception. | | |

| Analyte Matrix | Method Type | Method Number* | Latest JAOAC Reference |
|---|---|---|---|
| | ***Bioassy Methods*** | | |
| **Thiamine** | Growth | "938.12" | 25,0451 (1942) |
| | Rat growth method has same precision as chemical methods. | | |
| **Vitamin D** | | | |
| Milk | Rat | 936.14 | 46,0160 (1963) |
| | Bioassy is in terms of healing scores, and is not amenable to statistical analysis. Very few laboratories can perform this assay. | | |
| | ***Infant Formula*** | | |
| **Proximates** Mild-based | Grav./Titrimetric | 986.25 | 69,0777 (1986) |
| Elements | | | |
| CA,Cu, Fe,Mg, Mn | ICP Emission | 984.27 | 67,0985 (1984) |
| P, K, Na, Zn | Atom Absorp. | 985.35 | 68,0514 (1985) |
| Phosphorus Milk-based | Spectrophotometric | 986.24 | 69,0777 (1986) |
| Chloride Milk-based | Potentiometric | 986.24 | 69,0777 (1986) |
| Thiamine Milk-based | Fluorometric | 986.27 | 69,0777 (1986) |
| Riboflavin Milk-based | Fluorometric | 985.31 | 68,0514 (1985) |
| Vitamins A, D, and E Milk-based | Not recommended for approval | | |
| Vitamin B6 Milk-based | Microbiological/Turb. | 985.32 | 68,0514 (1985) |
| Vitamin C Milk- based | Titrimetric | 985.33 | 68,0514 (1985) |
| Niacin and Niacinamide Milk-based | Microbiological/Turb. | 985.34 | 68,0514 (1985) |
| Cobalamin (vitamin B12 activity) Milk-based | Microbiolgical/Turb. | 986.23 | 69,0777 (1986) |
| Pantothenic Acid Milk-based | Microbiological/Turb. | 43.025 [84] | 69,0777 (1986) |
| | Variability is high but acceptable (RSDR 20%; HORRAT 1.5); note adopted in expectation of obtaining "improved" method. Improved quality control required since some laboratories not aware they were producing outliers. | | |

The results from the method-performance studies for the proximates, inorganic nutrients, and water-soluble vitamins are acceptable, those for viatmins A and D are borderline, but those for vitamin E are unacceptable.

* = Surplus (No longer used)
SDU = Special Dietary Use
Grav. = Gravimetric
Titr. = Titrimetric
Turb. = Turbidimetric

# Effects of Nutrient Data Changes on Results of NFCS 1977-78 and NFCS 1987-88

**Patricia M. Guenther and Betty P. Perloff**
**USDA Human Nutrition Information Service**
**Hyattsville, Maryland**

The 1988 Bridging Study was conducted by the U.S. Department of Agriculture (USDA) to facilitate the comparison of the results of the 1977-78 Nationwide Food Consumption Survey (NFCS) and 1987-88 NFCS (Guenther and Perloff, 1990). The purpose was to determine whether differences between interview and review procedures, food coding and gram weight conversion procedures, and nutrient data bases used could result in differences in estimated food and nutrient intakes.

For example, a series of probing questions was added to assist respondents in recalling food items that were thought to be often forgotten, and the Food Instruction Booklet was expanded from 4 to 18 pages to capture more detailed food descriptions and more accurate quantities of foods eaten. USDA updated the food code manual and weight conversion factors, and coding procedures were partially automated. The nutrient data base was updated to reflect changes in the composition of foods and improvements in the quality of the food composition data.

A field experiment was designed to test the effects of these changes. A split-sample approach was used as subjects were randomly assigned to one of two treatment groups. Group A (N=348) was interviewed by 1987 procedures, and Group B (N=349) by 1977 procedures. Group A dietary recalls, collected by 1987 interview procedures, were reviewed and coded using the 1987 food codes and procedures, and nutrient intakes were calculated using the 1987 nutrient data base. Group B dietary recalls collected by 1977 interview procedures were reviewed by 1977 reviewprocedures and coded twice independently--once using 1987 codes and procedures and once using 1977 food codes and procedures. Nutrient intakes for respondents interviewed by the 1977 procedures were calculated four ways: (1) using the 1987 codes, weights, and nutrient data base, (2) using the 1977 food codes and the 1987 weights and nutrient data base, (3) using the 1977 codes and weights and the 1987 nutrient data base, and (4) using the 1977 codes, weights, and nutrient data base.

To evaluate the overall differences between the results obtained using 1977 and 1987 procedures, results from the two independent groups, A and B, were compared using two-sample multivariate t-tests. To evaluate the differences due to specific procedures, other than interview, results obtained from Group B, computed in the different ways, were compared using repeated measures techniques. When differences were found among the various 1977 results, paired t-tests were used to determine where the differences existed.

When the intakes of food energy and 14 nutrients by women in the two groups were considered jointly, the overall difference was significant (p .10). The differences found for iron, magnesium, and thiamin--the three nutrients with significant univariate differences--could not be attributed to differences in the interview or coding procedures. Rather, they were due to changes in the nutrient data base and the gram-weight conversion procedures.

For iron, the nutrient data base changes represented a combination of real changes fortification levels for some foods and improvements in the quality of the data for meat. For magnesium, the changes mostly reflected more recent but still limited data for coffee, while for thiamin the changes were mostly real.

For the most part, the effects of the various changes in survey procedures were slight and tended to offset each other. However, for four of the nutrients, we concluded that revising the 1977 estimated intakes was warranted becasue of improved data quality. Vitamins B-6 and B-12 and magnesium were revised because the data for them were more limited in 1977 than for the other nutrients reported (U.S. Department of Agriculture, 1984). We also revised the iron data because iron is a public health issue and because several important improvements had been made to the data base for this nutrient since 1977. This study has helped conceptualize the process for systematic tracking and revising of nutrient values for trend analysis.

## REFERENCES


Guenther, P.M., and Perloff, B.P. 1990. Effects of Procedural Differences Between 1977 and 1987 in the Nationwide Food Consumption Survey on Estimates of Food and Nutrient Intakes: Results of the USDA 1988 Bridging Study. Washington: U.S. Department of Agriculture, Human Nutrition Information Service. Nationwide Food Consumption Survey 1987-88, NFCS Rep. No. 87-M-1, 48 pp.

U.S. Department of Agriculture, Human Nutrition Information Service. 1984. Nutrient Intakes of Individuals in 48 States, Year 1977-78. Washington; Nationwide Food Consumption Survey 1977-78, NFCS Rep. No. I-2, 437 pp.

# Update of HNIS Survey Activities

**Alanna J. Moshfegh, M.S., R.D.**
**HNIS / USDA**
**Hyattsville, Maryland**

The Human Nutrition Information Service is responsible for conducting surveys to assess the American diet. In these surveys, we obtain three types of information relating to food consumption and dietary behavior: food used by households, food eaten by individuals, and assessment of dietary knowledge and attitudes of individuals.

USDA has been in the food consumption survey business for 55 years dating back to 1936 when the first survey measuring household food use was conducted. We began the individual food intake component of the survey in 1965. In 1985, we initiated continuous data collection with the Continuing Survey of Food Intakes by Individuals which includes only the collection of individuals intake information. In 1989, we introduced a new survey called the Diet and Health Knowledge Survey. It is being conducted as a telephone follow-up to selected respondents in the Continuing Survey. Today, survey work at HNIS is concentrating on surveys conducted from 1987 and on.

In the 1987-88 Nationwide Food Consumption Survey (NFCS), there is a large amount of data to process and analyze. Interviewers ask for information on the types, amounts, and costs of all foods used by the households during the previous 7 days. Then, the interviewer completes a 1-day recall of food intakes for each household member and instructs the respondents to complete a 2-day dietary record to be picked up by the interviewer 2-4 days later. In the 24-hour recall and 2-day record, we ask for information on the type of and amount of each food eaten, time and name of the eating occasion, and whether it was from home food supplies or obtained and eaten away from home. This information, on top of information collected on the individual's diet and health and sociodemographic characteristics such as education and income, is obviously a great deal of information to collect, and many people in 1987-88 were unwilling to provide it.

The household level response rate for the 1987-88 NFCS was 38 percent and the 1-day individual response rate was 31 percent, much lower than in previous surveys, and much lower than we are comfortable with. We believe one of the main reasons for the low response rate is due to the increased value and protection each of us places on our time. There is an increased proportion of women in the work force and therefore less likely to be home or to be willing to devote time to a long interview, increased number of surveys by many types of organizations, and greater concern about letting strangers into your home.

We have little doubt, however, that there was a problem with the heavy respondent burden of the survey. The average length of the interview was 2.7 hours, and many people refused to participate after being informed of the 2 requirements of the survey. One -- although not the only -- cause of this burden was conducting the household and individual components of the survey together.

Because of the urgent need for the NFCS data, the individual intake data was released on data tape last October before the impact of nonresponse was fully investigated. The tape included a warning notice describing the nonresponse rate and the potential for nonresponse bias.

Since last fall, we have conducted a series of studies on comparing the NFCS sample characteristics and data with other surveys including the March 1987 Current Population Survey (CPS), the 1987 National Health Interview Survey conducted by the National Center for Health Statistics, and comparisons with our previous surveys -- the 1977-78 Nationwide Food Consumption Survey and the Continuing Surveys. We also contracted for an independent expert panel to assess causes and implications of nonresponse.

The first study conducted was the comparison of the unweighted NFCS sample with estimates of the distribution of sociodemographic characteristics derived from the March 1987 Current Population Survey that are related to dietary intake. For most of them, the NFCS sample represented the population fairly well. One area of difference was that the NFCS included fewer female heads of households who were employed. Other small differences were that the NFCS sample has a slightly higher proportion of households with lower incomes and fewer 15- to 24-year-olds. When our weighting factors, developed by Wayne Fuller, Ph.D. at Iowa State University, are used with the sample data, they yield estimates that match the CPS estimates for these characteristics. The weighting scheme also corrects for deviations from the survey design related to uneven spacing of interviews over the calendar months of the year and over the 7 days of the week. This is expected to reduce the magnitude of nonresponse bias in the estimates; however, the use of so many control factors has the potential to increase the variances of the estimates.

Second, we compared estimates from the NFCS with estimates from the 1987 National Health Interview Survey (NHIS), conducted by the National Center for Health Statistics. The 1987 NHIS, with 22,080 respondents, had a response rate of about 95 percent. Therefore, nonresponse bias is considerably less likely in that survey. Some questions on sociodemographic and health variables were asked in an identical, or very similar, manner in NFCS and NHIS, making possible some direct comparisons of these characteristics. The estimated mean levels of several characteristics were in fairly close agreement between the two surveys. These characteristics included self-reported weight and height, household income expressed as a percent of the federal poverty level; education of the household head; employment status; age; and health status.

For other characteristics, there was less agreement between the estimated mean levels from the two surveys. For example, degree of urbanization, race, ethnic origin, smoking status, and use of vitamin and mineral supplements.

We turned also to comparisons with our own previous surveys. We compared 1987-88 NFCS estimates of energy and nutrient intakes with those from the NFCS conducted 10 years earlier. We also examined estimates from the Continuing Survey of Food Intakes by Individuals (CSFII), conducted in 1985 and in 1986.

These surveys had different methodology, design, and target sample; but allfour surveys included a 24-hour recall of dietary intake in April and May, and comparisons were limited to those data. While we were looking for evidence of nonresponse bias in NFCS, it appears that what we found were differences between NFCS and the Continuing Surveys. The estimates from the two NFCS were generally more similar to each other than they were to estimates from the twoCSFII; and estimates of the intake of calories and nutrients were generally lower than in the CSFII. These estimates most likely reflect the differences in respondent burden between the NFCS and the CSFII.

And finally, we contracted with the Life Sciences Research office of the Federation of American Societies of Experimental Biology to convene a 3-member independent expert panel of statisticians to review the 1987-88 NFCS survey design and execution as well as the nonresponse studies conducted by HNIS. Based upon our studies and their report, it clearly cannot be demonstrated nor can it be stated that there is or is not nonresponse bias in the NFCS. The NFCS data has serious potential for error, and error has two components -- bias and variance. We believe the potential for bias has been limited to the best of our ability with our weighting procedures. As for variance, the potential for it increased by weighting; and users can and should objectively evaluate variance for estimates of interest by using software appropriate for multi-stage survey data, such as PCCARP, SESUDAAN, and OSIRIS, to calculate statistics such as standard errors and coefficients of variation. We are not stating that the data cannot be used. However, we are recommending that users of these data carefully balance their need and the tolerance for error in their specific application against the potential for nonresponse bias in the 1987-88 NFCS dataset. Whenever possible, confirmatory data from other sources should be sought to support results based on analysis of these data.

A report of the nonresponse investigations is in preparation. It will include the report of the Expert Panel. Also, with all data released on the NFCS, a three-page statement on nonresponse issues including response rate, weighting factors and cautions regarding increased variance and nonresponse bias is being included.

Reporting plans for the NFCS data have been limited because of the nonresponse. Data tapes of the individual component were released last October and the household component should be available by this fall. A methodology report on the 1988 Bridging Study, was published. Two statistical data reports will be published -- one on food consumption and dietary levels of households and one on food and nutrient intakes by individuals.

The Continuing Survey of Food Intakes by Individuals was initiated in 1985 to provide continuous monitoring of the dietary status of the Population. The survey was conducted again in 1986, but discontinued during 1987 and 1988 while the NFCS was in process. In 1989, the Continuing Survey began again and we are now in the third year of data collection.

The sample is smaller than that for the NFCS. For each year, the total sample is 2,250 households including both all income and low income households. By combining the data for several years, however, we will be able to provide information for a much larger sample. The dietary intake methodology for the current Continuing Survey is a 1-day recall followed by a 2-day record.

The response rates for the Continuing Survey are better than for the NFCS, but they are still not what we would like. The overall 1-day response rate for the 1989 survey was 57 percent for the basic survey and 63 percent for the low income survey. The agency is taking steps to try to improve these rates and to lower the respondent burden.

How accurate are people's perceptions about their diets? This is one question we hope to be able to answer with our Diet and Health Knowledge Survey. The DHKS is a telephone follow-up to the CSFII. About six weeks after participating in the CSFII, the person in the household identified as the main meal-planner/preparer was recontacted by telephone and asked a series of about 36 questions.

Conceptually, the DHKS probes issues relating to the Dietary Guidelines for Americans. Survey questions are generally of five types: Attitudes about one's own diet; knowledge about foods; food preparation practices followed; food shopping practices including use of labels; and food safety concerns. Next year, a report will be issued on nutrition attitudes and dietary status covering information from both the CSFII and DHKS. The report will show relationships between knowledge and attitude parameters, health-related behaviors, and the dietary status of the main meal-planner/preparer.

Where do we go from here? Commitment of the Department of Agriculture to nutrition monitoring and food consumption surveys is strong as evidenced by our increasing number and types of surveys particularly during the 1980's. Because of the low response rates in the NFCS and in light of the passage ofnutrition monitoring legislation last year, the Human Nutrition Information Service believes that now is the time to begin an in-depth review of the Agency's survey activities. To take a step back and plan for the 1990's and into the next century to continue improving and enhancing food consumption surveys conducted by USDA.

To begin, two interagency agreements have been signed with the U. S. Bureau of the Census. Staff of the Census Bureau have been meeting with HNIS staff on a regular basis to get acquainted with the survey methodology and survey operations. The agreement covers numerous areas including:

- o Assisting us in planning future surveys including the ongoing CSFII and the next NFCS;
- o Providing an in-depth review of the survey questionnaire including suggestions for reducing respondent burden and cognitive testing of questions to be sure they measure what we think they measure;
- o Suggesting ways to improve survey management and build in better quality controls;
- o Providing on-going support during data collection and processing and suggesting ways to improve efficiency;
- o Reviewing reasons for nonresponse and making suggestions for needed changes; and
- o Conducting or advising on research to improve dietary intake methodologies. Some of this research will be directed to issues such as use of proxy respondents and the number and spacing of interviews to best measure dietary intake.

A second area of change is our new Survey Data Management System. This system has two purposes: (1) to provide an efficient state-of-the-art system for future HNIS contractors to use for coding dietary recall and food intake records, and (2) to provide automated procedures for efficient maintenance of the food code system and to monitor and measure quality of a contractor's performance. The centerpiece of this system is the Food Intake Analysis System, an interactive, user-friendly nutritional analysis software package, developed jointly by HNIS and the University of Texas School of Public Health.

A third area of focus is on contract management. We have in place a full-time contract manager and are providing staff with additional training in contract management. In addition, we are making weekly trips to the contractor's central offices to monitor survey operations.

In addition to the survey activities, HNIS, being the lead agency in USDA for nutrition monitoring, is busy implementing the requirements of the National Nutrition Monitoring and Related Research Act signed into law last October. We are working jointly with our colleagues in the Department of Health and Human Services in this effort. We are to establish an Interagency Board for Nutrition Monitoring. The Board has been established as of March 1991 by incorporating the functions and members of the Interagency Committee on Nutrition Monitoring. This Board is co-chaired by the USDA Assistant Secretary for Food and Consumer Services, Catherine Bertini, and the DHHS Assistant Secretary for Health, Dr. James Mason.

This Board is to assist in implementing the coordinated nutrition monitoring program. We are to establish a 9-member Nutrition Monitoring Advisory Council made up of experts in nutrition monitoring from outside the Federal government with 5 members appointed by the President and 4 by Congress. The Advisory council was established by Presidential Executive Order on January 25, 1991. Recommendations for the President's 5 appointments have been jointly made by USDA and DHHS and sent to the White House. There are also a number of reporting requirements including an annual interagency budget on nutrition monitoring to be reported to Congress.

The last requirement of the law I want to discuss is the development of a 10-Year Comprehensive Plan for National Nutrition Monitoring. The draft plan is to cover 10 years and address how the Federal government is to implement a comprehensive and coordinated program in nutrition monitoring. A draft of the plan is to be published in the Federal Register for public comment in late October 1991. There will be a 90-day comment period with a final plan to be published in the spring of 1992. We are working to solicit input from users of the monitoring data. Many sessions are planned at professional meetings to solicit input and recommendations. Major sessions are planned at the American Dietetic Association's and the American Public Health Association's annual meetings this fall. Your input as users of the data and information from the monitoring system is critical to development of an effective plan and program. Please let us hear from you.

In closing, let me leave you with an invitation to the HNIS Resources Conference on Nutrition Monitoring that we are holding in cooperation with theUniversity of Texas School of Public November 7-8, 1991 in Bethesda, Maryland. The conference will describe nutrition monitoring activities of HNIS includingour food consumption surveys, the nutrient data bank and other programs. Registration is free but limited due to meeting room space. A special feature is that all attendees will receive a data set of the 1989 Diet and Health Knowledge Survey on diskette. We hope you will join us in November 1991.

# USDA Nutrient Data Base Update

**Betty Perloff**
**HNIS / USDA**
**Hyattsville, Maryland**

The Human Nutrition Information Service (HNIS) is responsible for the preparation of food composition tables and computerized data bases representing foods consumed in the United States. Agriculture Handbook No. 8 (AH-8) and its corresponding computer data set, the USDA Nutrient Data Base for Standard Reference, represent one of the largest compilations of food composition values in the world--a compilation that must continue to expand and improve to meet the challenges presented by nutrition monitoring and foods research. This paper provides information about nutrient data releases and related activities at HNIS since we reported last year at the Fifteenth National Nutrient Data Bank Conference.

### Nutrient Data Bank Bulletin Board

Last year at the Nutrient Data Bank Conference we announced the installation of an electronic bulletin board, and this year we are pleased to announce its success. Its purpose is two-fold. First, it provides a ready source of information about nutrient data releases, upcoming Nutrient Data Bank Conferences, and relevant publications and activities. Second, it serves as a mechanism for making small nutrient data and other computer files available quickly and inexpensively.

These small nutrient data files include updates to the Standard Reference Data Base as they become available. For example, when the revised Agriculture Handbook No. 8 section on beef was released and the Standard Reference Data Base updated last year with those values, the newer values for beef were placed on the bulletin board. Instructions were included for updating the Standard Reference Data Base by replacing the older beef values with the new values.

### Agriculture Handbook No. 8

Since the meeting last year, we have published the 1990 Supplement to Agriculture Handbook No. 8. The series of annual supplements began in 1989 to make selected updates to foods within food groups that do not need an overall revision. It contains replacement pages for foods that have had data updates and insertion pages for foods that are being added to the handbook. The 1990 Supplement includes information for 10 different food groups. The 1991 Supplement is in preparation and will include revisions to fresh pork items, reflecting leaner pork now on the market.

The complete section on Snacks and Sweets, AH-8-19, is currently being prepared for printing and is expected off press this fall. It contains approximately 300 items.

The section on Baked Products, AH-8-18, which contains over 400 items, will be sent to food companies who submitted data in July. Companies are asked to review data for their products and to identify those food items that have changed recently. This type of review usually results in a few updates to values even before they are published. The Baked Products section has been one of the most complicated to complete in the entire AH-8 series because of the complexities of the foods and also because of the large number of foods involved.

### USDA Nutrient Data Base for Standard Reference

The computerized data set corresponding to the AH-8 series is the USDA Nutrient Data Base for Standard Reference. It includes data for all food groups and is updated approximately once a year to reflect any updates that have been published since the previous release. The current version, release number 9, became public in 1990. It contains the latest published data for each food group, but does not yet include the data from the 1990 Supplement. Release 10 will be available either later this year or early in 1992. It will reflect updates from the 1990 Supplement, the 1991 Supplement, and AH-8-19 on Snacks and Sweets.

We are very pleased to announce that since last year our systems staff have begun development of three computer programs for use with the Standard Reference Nutrient Data Base. The first program, "Standard Reference--Nutrient Analysis Tool (SR-NAT)," is a nutrient analysis program using a menu system for food identification and selection. SR-NAT uses the abbreviated version of the data base for microcomputers as the food and nutrient data base. This abbreviated version contains 21 specific nutrients.

The second program is called "SR-NAT plus." SR-NAT plus uses the full version of the data base for microcomputers, allowing the user to select up to 30 of the 81 nutrients that are present.

The third program, "Leveler," is a utility program for converting the Standard Reference data base into a format that can be used with data base management programs. Eventually, these programs will be placed on our bulletin board.

### Survey Nutrient Data Base

The Survey Nutrient Data Base is maintained especially for analysis of nationwide dietary intake survey data and is used not only by USDA but also by the Department of Health and Human Service's National Center for Health Statistics (NCHS) for the National and Hispanic Health and Nutrition Examination Surveys. Different releases of the data base, specific for the time period, are used for the different surveys conducted by HNIS and NCHS. A new release of this data base is awaiting final clearance and should be available shortly.

Many of the values on this data base have been calculated based on recipes for food mixtures. Food mixtures containing fat as an ingredient and fried foods have been calculated in several ways. First, they are calculated using a fat considered common for the specific food item. This set of values is referenced as the default values; they are used when survey respondents cannot designate the type of fat used in food preparation. Those mixtures are also calculated several additional times using different types of fats, and those resulting values are used when respondents can be specific about the fat used in food preparation.

Likewise, items with salt as an ingredient are calculated both with and without the salt. The values containing salt are considered the default values. All sets of values from these various calculations are placed on this data base and are included on the version released to the public through the National Technical Information Service.

A special version of the Survey Nutrient Data Base has been created for the bulletin board. First, it includes only the default values for items containing fat as a component. For example, the data for "green beans, cooked with fat added," represent green beans with margarine added. Second, it includes two sodium values for some items. For items with salt as a component, one sodium value will represent the default value. The second sodium value will represent the recipe caluculated by omitting the salt ingredient. For example, the recipe for green beans with fat added includes salt:

Green beans, cooked
Margarine
Salt

The two sodium values are:

Sodium - default = 253 mg/100 g
Sodium - salt omitted from recipe = 36 mg/100 g

### Food Intake Analysis System

A food intake analysis system using the Survey Nutrient Data Base has been developed jointly by HNIS and the University of Texas School of Public Health. This system also includes and gives users access to several other data files that form the technical support system for the survey. These include the Manual of Food Codes and Descriptions for Individual Intakes, the Primary Data Set used for basic ingredient data in the recipe calculations, the Primary Data Set description file, a file of nutrient retention factors, and the recipes used by HNIS for calculating food mixtures. Several of these files were enhanced by graduate students assigned to this project to make the files easier to use. For example, the descriptions used in the various files were tested for comparability and made consistent from one file to another. In addition, a new file was created to represent moisture and fat changes during cooking based on data from our recipe file and Agriculture Handbook 102, "Food Yields at Different Stages of Preparation." The moisture and fat changes are accessed and used by the recipe calculation feature of the system. This system is ready and awaiting our final release of the Survey Nutrient Data Base.

### Food Composition Data Working Group

Several food composition activities related to nutrition monitoring are under way and will effect future food composition research and data bases. In 1989 the Food Composition Data Working Group was formed by the Interagency Committee on Nutrition Monitoring, since renamed the Interagency Board for Nutrition Monitoring and Related Research. The working group's purpose is to identify food composition data needs for the National Nutrition Monitoring System, to propose options and priorities for improving the utility of food composition data for nutrition monitoring and related research applications, and to facilitate coordination among member agencies in the area of food composition measurement and research. Members of the working group are:

Betty Perloff, Human Nutrition Information Service, Co-Chair
Gary Beecher, Agricultural Research Service, Co-Chair
Ruth Matthews, Human Nutrition Information Service
Ronette Briefel, National Center for Health Statistics
Margaret McDowell, National Center for Health Statistics
Jean Pennington, Food and Drug Administration
John Vanderveen, Food and Drug Administration
Betsy Frasao, Economic Research Service
Susan Pilch, National Institutes of Health

The working group has identified several issues needing attention and specified activities to be undertaken to address those issues:

- *Develop criteria for establishing overall quality of a data base for priority nutrients.*

  The working group has identified four criteria to be factored into an evaluation scheme: (1) adequacy of analytical method for various categories of foods at different levels; (2) representativeness of the data base with respect to numbers of samples upon which individual values are based; (3) representativeness of the data base with respect to amount of analytical versus estimated values; and (4) adequacy of data from standpoint of how recently the individual values have been updated or re-evaluated. Further development of these four criteria into a workable systematic model for data base evaluation is in various stages.

- *Using the criteria developed from the previous activity, evaluate quality of food composition data base for priority nutrients.*

- *Establish criteria for developing, documenting, and using a food composition data base for trend analysis.*

  This is an important issue because of the need to compare nutrient intake data over several years. As food composition values improve, adjustments may have to be made to nutrient intake data from previous years so that comparisons from one time period to another are comparable. For example, the cholesterol value for eggs was lowered in 1989 as a result of improved analytical methods. Previously, in 1985, the intake of cholesterol by women 19 to 50 years of age was 280 milligrams per day as reported in the Continuing Survey of Food Intake by Individuals. The newer cholesterol value for eggs would lower that estimate by about 9 percent, and an adjustment should be made when future cholesterol intake estimates are compared with the 1985 estimated intake.

  The working group has agreed upon a basic design for a data base for trend analysis that will be a modification of the current Survey Nutrient Data Base. It will contain multiple sets of nutrient values for foods when necessary to reflect food changes that have occurred. Included with each set of values will be a starting and ending date reflecting the effective time period covered by the values, with the earliest beginning time period set at 1985. For example, the iron content of Grape Nuts changed during 1988 from 9.53 mg/100 g to 28.6 mg/100 g. The data base for trend analysis will include both values with appropriate starting and ending dates for each. Values that change because of data improvements will be replaced by the new values. For example, the newer cholesterol values for eggs will replace the older value and be used retroactively. When it becomes necessary to recalculate previous years' nutrient intake estimates, values appropriate for the specific time period can be used.

  The system files that are used to create the data base, i.e., the primary data set and recipe file, will be coded to indicate whether changes occur because of changes to foods, such as the iron in the breakfast cereal, or because of improvements in data, such as the cholesterol in eggs.

- *Establish criteria which constitute satisfactory documentation of food composition data base.*

  Documentation requirements are being considered currently by the working group, and a special session to receive input from data base users about their documentation needs is being held during this conference. Written comments, are also welcomed and may be sent directly to me at HNIS.

- *Identify foods and other items for which more information is needed, i.e., ethnic foods, recipes, retention and yield factors.*

- *Identify the needed specificity of food item descriptions, e.g., homemade versus commercial foods.*

- o *Identify needed improvements of measurement systems for nutrients in foods, i.e., methods, quality control materials, etc.*
- o *Develop criteria for estimating data (imputing) for foods for which there are no analytical values.*

**Summary**

In summary, work continues at HNIS on the continual revision of Agriculture Handbook No. 8, the Nutrient Data Base for Standard Reference, and the Survey Nutrient Data Base, as well as research activities to strengthen and expand the data base. Three new utility programs for use with the Nutrient Data Base for Standard Reference are being developed, and food composition activities related to the nutrition monitoring legislation are expanding. We look forward to increasing the utility of our Nutrient Data Bank Bulletin Board and have added several items at the request of users. Comments and suggestions for addition of new items or information are always welcome.

# NHANES III Update

**Robert J. Kuczmarksi, Ph.D., R.D.**
**Margaret McDowell, M.P.H., R.D.**
**National Center for Health Statistics**
**Hyattsville, Maryland**

## INTRODUCTION

The National Center for Health Statistics (NCHS) is an agency of the U.S. Public Health Service and is one of the Centers of the Centers for Disease Control (CDC). The primary mission of the NCHS is the collection and dissemination of health statistics for the nation. The NCHS produces a wide range of data that are used for health research, administration, planning, and education. There are a number of data collection systems within the NCHS. The National Health and Nutrition Examination Survey (NHANES) is only one of the data collection systems, but is unique in that it collects data on the health and nutritional status of Americans through interviews and direct physical examinations.

Operating from a Mobile Examination Center (MEC), a set of four interconnected and specially designed and equipped trailers, the HANES staff travels across the country to reach people selected to participate in the survey, and they administer standardized examination and laboratory tests. Before visiting a community, the examination team is preceeded by an interview team that visits individual households to select sample persons as part of the complex sample survey design for this nationally representative survey. The interviewers administer health and nutrition questionnaires to the subjects prior to inviting them to the MEC, which is subsequently brought into the local community.

## BACKGROUND

As a brief background, from 1959-84, the NCHS conducted six separate health examination surveys (Table 1). The first National Health Examination Survey (NHES I) focused on the prevalence of selected chronic diseases in civilian, noninstitutionalized adults. The NHES II and NHES III were devoted to growth and development of children and adolescents. In 1969, the Department of Health, Education, and Welfare established within the NCHS a continuing activity to measure the nutritional status of the U.S. population, and to monitor changes in nutritional status over time. An NCHS task force decided to combine the nutrition component with the NHES to permit relating nutritional variables to health measures that were already being collected. Thus, the first National Health and Nutrition Examination Survey (NHANES) was created. The most recently completed, nationally representative NHANES was conducted from 1976-80. The Hispanic HANES (HHANES), conducted from 1982-84, was a special survey of three selected subgroups of the population in selected areas of the U.S. rather than a national probability sample. The NHANES III began data collection in October 1988 and will continue until the fall of 1994. Persons are included in the NHANES III beginning at two months of age and for the first time in the NHANES, there is no upper age limit for participation in the survey.

Table 1. **National Health and Nutrition Examination Surveys.**

| Surveys | Dates | Ages |
|---|---|---|
| NHES I | 1960-62 | 18-79 years |
| NHES II | 1963-65 | 6-11 years |
| NHES III | 1966-70 | 12-17 years |
| NHANES I | 1971-75 | 1-74 years |
| NHANES II | 1976-80 | 6 mo.-74 years |
| HHANES | 1982-84 | 6 mo.-74 years |
| NHANES I Follow-up | 1982- | 25-74 years |
| NHANES III | 1988-94 | 2 mo.+ |

## NHANES III CHARACTERISTICS

### *Sampling Issues*

The sample for the NHANES III consists of large samples of Black, Hispanic, and other groups. Approximately 30 percent of the sample will be Black, 30 percent Mexican American, and 40 percent of the sample will include White presons among other subgroups of the U.S.population. The sampling scheme consists of 81 counties in 26 states for a total of 88 locations that will be visited. The survey consists of two National samples that will be conducted in successive 3-year cycles. The first cycle will be completed in the fall of 1991 and the second cycle will run from 1991 through 1994. Approximately 40,000 persons will be interviewed and approximately 30,000 persons will be examined. While the NHANES III continues to exclude military and Native American reservations, nursing homes, long-term care hospitals, and prisons, it is of interest to note that Hawaii and Alaska are included in the sampling frame. Hawaii was included in the NHANES II (1), and Alaska will be visited in the second phase of the NHANES III.

### *Structure*

The NHANES III household interview consists of four parts (Table 2) and it is here that detailed questionnaires obtain information on items such as family relationships, basic demographic information, participation in income assistance programs, and health status. Relevant to the nutrition component, participation in the Food Stamp Program and the WIC Program is assessed, as are the use of vitamin and mineral supplements. Infant feeding practices, such as breast feeding, the introduction of solid foods, and similar questions, are also asked. Interview questions assess alcohol intake in the past 12 months, the frequency of eating breakfast, quantity of tap water consumed in a day, type and frequency of salt use, availability of food or money to buy food, and a food frequency that is discussed in more detail below.

The focus of the examination component is on measures associated with chronic diseases, many of which are known to have relationships with diet and eventually can be analyzed in that connection. In addition, a comprehensive nutritional assessment is completed.

Table 2. **4-Part Household Interview**

1. Household screener questionnaire
2. Family questionnaire
3. Household questionnaire
   - Age 2 mo.-16 yrs.
   - Age 17+ yrs.
4. 3 seated blood pressures

### *Nutritional Assessment*

The nutritional assessment component includes an extensive array of anthropometric measurements that will be used primarily to assess underweight, overweight, and obesity and the inter-relationships among these conditions and a number of other risk factors and chronic conditions and diseases. Body measurements are also used to assess growth in children and adolescents. In fact, it is intended that measures from the NHANES III will be used to update the NCHS growth charts. The nutritional biochemistry measurements from the laboratory analyses use blood and urine specimens to assess an extensive list of lipids, vitamins, minerals, and other factors related to nutritional status (2).

### *Dietary Intake*

Dietary intake is a vital part of nutritional status assessment and was considered at length and very carefully during the planning phase that preceeded the NHANES III. In 1986, prior to the start of the survey, a workshop was held in which comments and opinions were considered from a wide range of experts from the nutrition community. It was decided that a single dietary methodology could not meet the needs of all the major nutritional objectives planned for the survey. A decision was made to administer at least one 24-hour dietary recall to provide detailed quantitative food and nutrient intake data for the U.S. population and selected subgroups of the population. In addition, it was decided that a qualitative food frequency instrument would be used to assess food items and food groups that were used in the previous 30 day period. For the purpose of monitoring trends over time, a food-frequency list that could be linked with data from the previous HANES surveys was developed. The food frequency targets dietary sources of nutrients that are believed to be associated with selected chronic conditions. Furthermore, foods reported by selected ethnic groups in previous HANES surveys were also included to make the instrument appropriate for population subgroups. As shown in Table 3, the food frequency is administered in the MEC for persons aged 2-11 months and 12-16 years. Persons aged 1-11 years do not receive the food frequency questionnaire, and for sample persons aged 17 years or more, the food frequency is administered as part of the household interview. The latter is important in that it will yield dietary intake data for the subgroup of persons who do not come to the MEC.

Table 3. **Administration of the NHANES III Food Frequency.**

| Age group | Place | Respondent |
|---|---|---|
| 1-11 mo. | MEC | Proxy |
| 1-11 yr. | MEC | ----- |
| 12-16 yr. | MEC | Self |
| 17+ yr. | Household | Self |

***24-hour dietary recall***

A 24-hour recall is collected in the MEC for all examined persons using an automated dietary data collection and interview system that was developed with support from several government agencies by the NCHS and the Nutrition Coordinating Center (NCC) at the University of Minnesota. This data collection system, known as the NHANES III Dietary Data Collection (DDC) System was described in detail at the 1989 National Nutrient Data Bank Conference held in Iowa City, Iowa and was published in the proceedings of that meeting and elsewhere in the scientific literature (3-4). For each sample person, a trained bilingual dietary interviewer records the quantity of every item of food or drink consumed during the 24-hour period prior to midnight before the day of the interview, enabling estimates to be made of macro- and micronutrients and energy. The dietary interviewers conduct open-ended interviews using structured probes to ensure standardized data collection. Data are collected on brand name products, ingredients, cooking methods, and the use of fat and sodium in food preparation. Information is also collected on the time of day the food was consumed, the name of the meal or snack, and the place where the food was consumed. In addition to the characteristics of the DDC shown in Table 4, abstract geometric food models and household measures along with the ability to edit dietary recalls both during and after the dietary interview, and the ability to record information during the actual interview about foods that are not currently in the system, constitute part of a system that provides, as one of its most important features, automatic coding of foods to the USDA database.

Table 4. **Automated 24-hour dietary recall features**

-- Detailed description of foods/beverages consumed
-- Portion sizes/use of food models
-- Time of day/name of meal, snack, etc.
-- "Where consumed"
-- Salt/fat added in preparation
-- Links foods eaten together as a multicomponent food

At this time, data from the NHANES III, including the 24-hour recall data, are not yet available for analysis and it is unlikely that the first results will be available for review prior to the fall of 1992 when the data tape for the sampling weights of the sites included in the first half of the survey is scheduled to be completed. The only statistics that are currently available for any of the components are on the response rates of survey participants. Response rates for the dietary recall component are high. Approximately 98 percent of the examinees completed the dietary interview component of the MEC exam. For the first 36 survey sites of the NHANES III, 12,435 24-hour recalls have been completed. The response rates by age are shown in Table 5. From this table, it can be seen that a small decrease in the response rate appears to be associated with increasing age.

As shown in Table 6, the 24-hour recall is administered for the youngest sample persons through proxy respondents, usually the mother or another caretaker who is familiar with what the child has eaten. As mentioned, all sample persons who visit the MEC are asked to complete the dietary recall.

***Telephone Recall***

In addition to the single 24-hour dietary recall obtained in the MEC, supplementary funding has been received from the National Institute on Aging to conduct a special dietary study on persons aged 50 years

Table 5. **24-hour dietary recall response rates by age**

| Age group | # completed | % completed |
|---|---|---|
| 2 mos.-5 yrs. | 3316 | 98.9 |
| 6-11 yrs. | 1375 | 98.6 |
| 12-19 yrs. | 1235 | 97.7 |
| 20-39 yrs. | 2641 | 97.9 |
| 40-59 yrs. | 1739 | 97.2 |
| 60-74 yrs. | 1341 | 96.5 |
| 75 + yrs. | 788 | 95.4 |

Source: NCHS, NHANES III (Stands 1-36)

Table 6. **Administration of the NHANES III 24-hour dietary recall**

| Age group | Place | Respondent |
|---|---|---|
| 2 mos.-5 yrs. | MEC | Proxy |
| 6-11 yrs. | MEC | Proxy/self |
| 12-49 yrs. | MEC | Self |
| 50 + yrs. | MEC | Self |
| | Telephone | Self |

and older. Two additional 24-hour recalls are collected by telephone at 8 and 16 months after the itial MEC visit. This initiative is known as the Supplementary Nutrition Survey of Older Americans (SNS). Data collection is performed under contract by the Westat Corporation. Among the major objectives for this special study is the collection of additional days of dietary intake data to improve the quantitative estimates of the usual intakes of the target population.

The same NHANES III DDC interview that is used in the MEC is administered by trained, bilingual telephone interviewers. A food model booklet containing 2-dimensional drawings of the abstract food models and measurement aids used in the MEC is mailed to SNS respondents prior to each contact for use during the telephone interview. The average 24-hour telephone recall is completed in approximately 20 to 30 minutes, similar to the time required for the MEC in-person interview. Interview questions on self-reported health status, food sufficiency, special diets, and other items that were asked in the initial interview are repeated in the telephone interview.

Response rates for the first and second telephone 24-hour recall in the NHANES III SNS are shown in Table 7. As of March 1991, participants from the the first 18 survey sites had been recontacted for at least the first telephone interview at eight months after the MEC visit. From the 1,479 persons eligible for this interview, recalls were completed by 1,152 sample persons, indicating a response rate of 78%. Among the 10 stands for which the first and second telephone recalls were conducted at 8 and 16 months respectively, the response rate for the first interview was 77%, and for the second interview, 70%. Examples of reasons for nonresponse include not having a telephone, subject deceased, subject refusal, unable to contact subject or subject away from residence for prolonged period of time, and inability to locate proxy respondents. Data from the SNS will be processed along with the rest of the NHANES III dietary data sets.

Table 7. **Response rates at 8 and 16 months for the SNS**

| | Number | Percent |
|---|---|---|
| Stands 1-18 | | |
| Eligible (8 mo.) | 1479 | |
| Completed (8 mo.) | 1152 | 78 |
| Stands 1-10 | | |
| Completed (8 mo.) | | 77 |
| Completed (16 mo.) | | 70 |

Source: NCHS, NHANES III

### *Special Features of the NHANES III*

The telephone data collection is one of a number of special features of the NHANES III. Some of other features that are new to the NHANES III are also noteworthy. For the first time, the NHANES has a longitudinal component built into it to provide a better understanding of disease etiology and the natural history of disease. Present efforts focus on tracking current addresses of the sample persons by mail and reviewing the NCHS National Death Index, as the primary means of mortality follow-up. In the future, contingent upon appropriate funding, it is anticipated that sample persons may be reinterviewed by telephone and may even receive a modified home examination at a later time.

The precedent for a home examination has been developed and implemented as a part of the NHANES III. Persons who cannot or will not come to the MEC for a complete examination are offered the option of having a health technician visit them in their place of residence. This modified home examination is designed to improve the response rates on selected components of the survey, especially for the very young and the very old, and includes anthropometry, spirometry, cognitive and physical function, and a venipuncture (5). The home exam has helped to improve the overall response rates as shown in Table 8. To date, for the first 36 sites that have been surveyed, approximately 86 percent of the persons selected to participate in the survey have been successfully interviewed in their households. Those sample persons were then invited to the MEC for an examination and as indicated, 77 percent of those originially selected to participate have been examined in the MEC. The addition of the modified home examination has increased the examination response rate by approximately 1.3 percent, with 203 persons having been examined in their homes. When compared with the NHANES II and the HHANES, there has been an

Table 8. **HANES interview and exam rates**

| Survey | Number selected | Inter-viewed percent | MEC Exam percent | MEC+ Home Exam percent |
|---|---|---|---|---|
| NHANES II | 27,805 | 91.0 | 73.1 | |
| HHANES | 15,931 | 86.3 | 73.3 | |
| NHANES III* | 16,572 | 86.2 | 77.0 | 78.3 |

Source: NCHS, (*Stands 1-36)

increase in the overall response rate of approximately 4 percent, even when the home examination is not taken into account. Table 9 shows the percentage of persons examined thus far in each age and sex group. From this table, it becomes apparent that the MEC response rates decrease progressively with age, and that for the hardest to reach age group (75 + years), the home exam approach has increased the response rate by 8.1 percent.

Table 9. **NHANES III examination response rates**

| | Male | | Female | |
|---|---|---|---|---|
| | MEC exam | MEC + home | MEC exam | MEC + home |
| 2 mos.-5 yrs. | 86.5 | 87.9 | 88.8 | 89.6 |
| 6-19 yrs. | 87.6 | 87.7 | 84.2 | 84.2 |
| 20-44 yrs. | 74.5 | 74.7 | 80.7 | 80.9 |
| 45-59 yrs. | 69.5 | 70.1 | 72.8 | 73.3 |
| 60-74 yrs. | 69.3 | 71.4 | 65.3 | 67.5 |
| 75 + yrs. | 58.8 | 66.9 | 53.7 | 61.8 |

Source: NCHS, NHANES III (Stands 1-36)

Another special feature of the NHANES III is the complete automation of the MEC which will help to facilitate more accurate data collection and more rapid data release. In addition to the microcomputer that runs the automated dietary recall, we have several other microcomputers on board, as well as a VAX minicomputer. Data collection in the MEC is completed online with daily back-ups, edit checks, and built-in quality control procedures. Data are sent to the NCHS and to contractors in machine readable format. This approach allows for built-in edit checks and assists in controlling the quality of the data.

The NHANES III was designed to have both a descriptive and an analytic orientation. With regard to nutrient intakes, descriptive data will be available for population subgroups by age, race, and sex, and other demographic variables as in previous HANES surveys. In addition, the analytic datasets will allow the study of associations among nutrient intakes and a variety of chronic conditions. The associations between nutrient intakes and clinical indicators of cardiovascular disease are prime candidates to be studied. Related to this, dietary data from the NHANES III are anxiously awaited to evaluate public education activities such as the National Cholesterol Education Program. For the first time, the NHANES is attempting to assess risk factors for osteoporosis of the hip by measuring bone density at the hip. The food frequency, 24-hour recall, and other specialized interview questions on the historical intake of dairy products at various stages of the life cycle will provide further insight into the risk factors associated with this condition. Beyond nutrition and chronic disease, data from the NHANES make a vital contribution to nutrition status monitoring, allowing the study of vitamin and mineral deficiencies and toxicities for example, as well as the impact of other dietary factors on health outcomes. For example, there will be data on growth in children and overweight and obesity in all age groups that may be analyzed in relation to the dietary intake of energy and selected macronutrients.

**Behind the Scenes**

Finally, for persons who have not had the opportunity to participate in the current NHANES III, a behind the scenes look at the major components included in this survey will provide an idea of the types of data that may be anticipated for future analysis in the connection between diet, nutrient intake, and health outcomes.

There is a physicians's examination in which, among other assessments, three seated blood pressure measurements are recorded. As indicated previously, the home interviewers are also trained to record 3 seated blood pressure measurements in the home for a total of six readings that can be averaged and analyzed. There is still interest in the connection between sodium as a nutrient and blood pressure levels. The automated dietary interview system does collect information on fat and sodium in food preparation, and this will be available in addition to the detailed data from the survey nutrient data bank on nutrients from the food items that are consumed.

A certified ultrasonographer scans the gallbladder for gallstones. A recent report has suggested an association between the increased consumption of total calories, fat, and sugar, as well as the decreased consumption of high-grain fiber, and gallbladder disease (6).

A dentist conducts a complete dental exam, looking for caries, periodontal status, tooth loss, soft tissue lesions, baby bottle caries in preschool children, and similar conditions.

The health technicians are licensed X-ray technicians and take x-rays of the hand/wrist and knee of older persons in the search for signs of rheumatoid arthritis and osteoarthritis of the joints.

Because heart disease is still a major source of morbidity and mortality in the United States, a complete electrcardiogram provides information on cardiovascular disease and cardiac irregulatities including silent myocardial infarctions.

An allergy test consists of the application of 12 standardized food, insect, animal, and mold allergens.

Diabetic retinopathy, characterized by hemorrhaging blood vessels in the retina of the eye, is the leading cause of new cases of blindness in adults. The fundus photography procedure provides a picture of the retina that can be interpreted for this condition.

With the oversampling and emphasis on elderly persons in the NHANES III, an addition to this survey is the physical function test for persons aged 60 years or older. Exercises such as the time it takes to accomplish the everyday task of opening a door lock or walk a short distance are quantified.

While a major use of the food and nutrient intake data will be, for example, to help measure progress toward the Year 2000 Health Objectives for the Nation (7), it is hoped that other researchers interested in exploring the connections between nutrient intakes and the health outcomes measured in the NHANES III will begin to formulate research hypotheses or at least remember the sorts of data that will eventually become available for analysis through the public release data tapes.

**References**

1. National Center for Health Statistics. Plan and operation of the Second National Health and Nutrition Examination Survey, 1976-80. Vital and Health Statistics. Series 1, No. 15. DHHS Pub. No. (PHS) 81-1317. Public Health Service. Washington. U.S. Government Printing Office, July 1981.

2. McQuillan GM, Gunter EW, Lannom L. Field issues for the plan and operation of the laboratory component of the Third National Health and Nutrition Examination Survey. J Nutr. 120:1446-1450, 1990.

3. Feskanich D, Sielaff BH, Chong K, Buzzard IM. Computerized collection and analysis of dietary intake information. Computer Methods and Programs in Biostatistics. 30:47-57,1989.

4. Feskanich D, Buzzard IM, Welch BT, et al. Comparison of a computerized and a manual method of food coding for nutrient intake studies. J Am Diet Assoc. 88:1263-1267, 1988.

5. Harris T, Woteki C, Briefel RR, Kleinman JC. NHANES III for older persons: nutrition content and methodological considerations. Am J Clin Nutr. 50:1145-1149, 1989.

6. Tamimi TM, Wosornu L, Al-Khozaim A, Abdul-Ghani A. Increased cholecystectomy rates in Saudi Arabia. Lancet. 336:1235-1237, 1990.

7. Department of Health and Human Services. Healthy People 2000: National health promotion and disease prevention objectives. DHHS Publication No. (PHS) 91-50212. U.S. Government Printing Office, 1991.

# Status of INFOODS Publications

**William M Rand***

**John C. Klensin+**

As of the time of the conference, there were four major INFOODS publications pending. They are:

***Compiling Data for Food Composition Data Bases.*** This book discusses issues in the actual compilation and assembly of nutrient composition tables and data bases, including ways of handling missing values. It is expected to be published early in the last quarter of this calendar year by UNU.

***INFOODS Food Composition Data Interchange Handbook.*** This book discusses the actual INFOODS inter-regional model and formats for interchange of food composition data and associated descriptive information. It is expected to be published early in the last quarter of this calendar year by UNU, roughly simultaneously with the above.

***INFOODS guidelines for describing foods.***
Published: Truswell AS, Bateson DJ, Madfiglio KC, Pennington JAT, Rand WM, Klensin JC. "INFOODS Guidelines for Describing Foods: A Systematic Approach to Describing Foods to Facilitate International Exchange of Food Composition Data". *J. Food Composition and Analysis* 4, 1 (March 1991), pp. 18-38.

***Guidelines to the Production, Management, and Use of Food Composition Data.***
To be published by Eurofoods, probably late 1991 or early 1992.

*Professor of Biostatistics, Department of Community Health, Tufts University School of Medicine. Formerly Executive Secretary, INFOODS.

+Director, INFOODS Secretariat, Massachusetts Institute of Technology.

# Guide for Nutrient Data Users

**Grace J. Petot**
**Case Western Reserve University**
**Cleveland, Ohio**

The increasing accessibility of computers to professionals in research, education, and health care facilities is resulting in many more users of nutrient data. However, documented studies have shown errors and inconsistencies in applications of computerized nutrient data bases. Many users of nutrient data bases are searching for information and guidance to help them use the data appropriately and accurately. Definitive guidelines for the use of nutrient data are needed since fewer users are acquiring copies of up-to-date printed tables. Data bank conference proceedings are helpful but are not in wide distribution.

HNIS,USDA has agreed to cooperate with me and Case Western Reserve University to produce such a guide for users of USDA nutrient data and data from other sources. It is planned that such a publication will be made available in machine readable form to purchasers of USDA files and in printed form to users.

An outline has been prepared and reviewed by selected data base users and data base system developers. I would like to review this outline with you today and will be pleased to receive any comments or suggestions. Betty Perloff and David Haytowitz of the Nurient Data Research Branch of USDA are working with me. We plan to have it completed for publication this fall.

## Outline

1. Introduction
   - A. Purpose
   - B. Organization of guide
   - C. Suggestions for using

2. Definitions
   - A. Food composition tables, nutrient data bases
   - B. Related terms (household measures, yields, equivalents, retention, flags, impute, missing values, etc.)
   - C. Nutrient data base systems

3. USDA and food composition tables
   - A. History
   - B. Acquisition of data
   - C. Nutrient Data Base for Standard Reference
     - 1) Relationship to Handbook 8
     - 2) Foods, food names
     - 3) Nutrients
     - 4) Formats
       - a) Correspondence with old Handbook 8 (1963)
       - b) Correspondence with Revised Handbook 8
       - c) Explanation of fields
     - 5) Updating, release schedules
     - 6) Forms, formats available
       - a) Machine readable
       - b) On line file transfer
   - D. Revised Handbook 8 - printed
   - E. Notes about USDA data
     - 1) Sources
     - 2) Representative values
     - 3) Reliability
     - 4) Notes on nutrients, foods
     - 5) Levels of accuracy, precision, imputing
   - F. Other USDA food composition tables, reports, provisional tables

4. Data from other sources
   - A. Analysis methodologies
   - B. Sources of variation
     - 1) Laboratory

2) Calculated
   a) Recipes
   b) Product formulas
   c) Documentation

C. Nutrient/food component relationships - mois ture, solids, etc.

D. Sources, report formats, documentation
   1) Food industry
   2) Research journal reports
   3) Nutrition labels, product labels
   4) Laboratory analysis
   5) Other nutrient data bases, published tables

E. Special problems with some nutrients, foods
   1) Analytical methods
   2) Bioavailability
   3) Distribution in food supply
   4) Timeliness and adequacy of data

5. Using food composition data on a computer
   A. Storing data
      1) Food names
      2) Food codes
      3) Quantity definition
         a) Unit conversions
         b) Household measure equivalents
         c) Conversions from food models
         d) Conversions from dimensions
         e) Frequency factors
   B. Accessing data
      1) Calculations, formulas
      2) Coding decisions, guidelines
      3) Recipes
         a) Yield factors
         b) Retention factors
         c) Fat, water losses or gains
   C. Reporting data
      1) Summaries, averages
      2) Precision, rounding, missing data
      3) Dietary evaluations - RDA, % energy, ratios, etc.
      4) Formats

6. Maintenance of nutrient data base
   A. Updating
   B. Adding data - foods and/or nutrients
   C. Documentation
   D. Quality control
   E. Archiving

7. Tools, references for using nutrient data

# Maintaining Time - Related Databases for Dietary Data Collection and Nutrient Calculation

**I. Marilyn Buzzard**
**University of Minnesota**
**Minneapolis, Minnesota**

## Introduction

The need for comparability of food and nutrient intakes over time is being increasingly recognized by policy makers at both the local and the national levels, as well as by researchers who conduct long-term dietary intervention trials and other nutrition related research. Time-related databases are required for monitoring progress toward meeting specific nutrition objectives such as the Year 2000 nutrition-related Objectives for the Nation (1). We need to be able to assess the changes in what Americans are eating without the confounding due to using different nutrient databases or different coding schemes for surveys conducted at different periods of time. Similarly, medical researchers must have confidence that changes in nutrient intakes represent actual changes in the subjects' food consumption rather than artifacts of using different versions of a nutrient database.

Investigation of long-term trends in food and nutrient intake requires the use of databases which reflect the constantly changing marketplace and changes in food composition. These databases must also be maintained in a manner that permits recalculation of nutrients at any point in time to take advantage of the availability of improved food composition data. Recalculation of previously collected food intake data also permits the investigator to calculate intakes for other food components that may have been more recently added to the nutrient database.

The Nutrition Coordinating Center (NCC) at the University of Minnesota has developed nutrient database maintenance procedures to accommodate the needs of long-term clinical trials and other nutrition research studies requiring data collection over a period of many years (2-4). These procedures have now been extended to the more complex databases required for interactive collection and automated coding of food intake data (5,6). The NCC has also developed customized database maintenance procedures to meet the special needs of the National Health and Nutrition Examination Survey (NHANES III) (7).

The purpose of this paper is to identify the major factors affecting the comparability of dietary data over time, to describe the NCC databases used for collection and nutrient calculation of food intake data, and to summarize the database maintenance procedures required for the assessment of dietary data over time. These procedures are necessary not only for valid data analysis for long-term studies, but also for comparison of food and nutrient intake data among studies conducted during different time periods.

## Time-related factors

There are three general categories of time-related factors pertaining to the long-term collection and nutrient calculation of food intake data. These include: 1) marketplace changes; 2) changes in food preparation practices; and 3) the availability of new or improved food composition data, including both nutrient values and non-nutrient data (such as food densities). Most of the changes in the first two categories reflect actual changes in foods consumed, whereas most of the changes in the third category are related to the availability of better nutrient and non-nutrient data, rather than to actual changes in what people are eating.

Examples of marketplace changes include new products, new serving sizes of existing products, reformulations of existing products, and discontinued products. Current trends toward the use of less sodium and fat in preparing foods are examples of changes in food preparation practices. New analyses of foods and the use of improved analytical methodologies result in the ongoing availability of new and improved data on food composition. Database maintenance procedures that permit comparison of dietary data over time must account for all of these types of changes.

### General approaches to nutrient database maintenance for providing long-term stability of nutrient calculations.

Several different approaches have been used by national nutrition surveys and medical research studies to provide long-term stability for nutrient calculations. One approach is to freeze the nutrient database at the beginning of a study and continue to use that database throughout the study. A number of clinical trials have used this approach in the past. The major problem with this approach is that the study is unable to take advantage of new and better data that become available over the course of the study.

A second approach is to calculate nutrients only at the end of a study, using the most current data available at that time. The benefit of this approach is that the studies are able to take advantage of new data that become available during the period of the study. The United States national nutrition surveys, including the USDA Nationwide Food Consumption Surveys and the DHHS National Health and Nutrition Examination Surveys, have used this approach in the past to provide stability during the period of the survey. The problem with this approach for survey use is that it does not permit comparison with previous surveys that have used different nutrient databases. Another major disadvantage is that nutrients can be calculated only at the end of the study. Therefore, this approach is not an option for most clinical trials and other studies that require ongoing monitoring of differences between study groups or compliance to a dietary intervention protocol.

A third approach is to use the most current nutrient database available whenever nutrient calculations are desired. This approach requires the recalculation of all previously processed data so that all nutrient intakes are calculated from the same nutrient database. With the progressive decline in computer calculation costs and the increasing use of microcomputers, this approach is becoming a viable option for many studies.

The NCC has developed databases and database maintenance procedures that permit studies to use any of the three approaches described above. The remainder of the paper will describe the two major NCC databases and the database maintenance procedures used to enhance comparability of food and nutrient intakes over time.

### Overview of NCC databases

The NCCcMinnesota Nutrition Data System (NDS) is an automated system for interactive data collection and nutrient calculation. NDS requires two NCC databases: the Food Description database (FDdb) and the Nutrient database (Ndb). The FDdb contains the information required for interactive collection and automated coding of food intake data, and the Ndb contains the nutrient values required for nutrient calculation.

Figure 1 shows the three major steps involved in the interactive collection and automated processing of dietary data and the relationships among these steps and the two NCC databases. Interactive data collection is the first of the three major steps involved in this process. Regardless of the method used for collecting the data (eg, 24-hour dietary recall, food records, or diet history), identification of the foods and amounts consumed is captured by the computer in language commonly used by subjects for describing foods. Coding consists of assigning one or more food codes to each food description. These codes are linked to entries in the Ndb. The coding step also involves converting amounts expressed in common household units or food specific units (such as piece, medium, slice, or package) into gram weights for nutrient calculation. The final step involves the calculation of nutrients using the Ndb and calculation software.

The FDdb contains more than ten times as many food descriptions as the Ndb. This is due not only to the fact that many foods in the FDdb are composites of a number of individual ingredients included in the Ndb, but also to the fact that foods having similar nutrient profiles are often grouped together in the Ndb. Thus, the food descriptions listed in the Ndb are primarily generic food descriptions, any one of which may represent many different brand name products or food types. To permit maximum flexibility and specificity for describing foods, all non-nutrient information (such as recipes, densities, yield factors, and other amount conversion data) is maintained for each of the many foods in the FDdb, rather than for the few foods in the Ndb. The Ndb, which is the least complex of the two databases, has been maintained as a time-related database for many years. Therefore, I will address the maintenance of this database first.

### Specific guidelines for maintaining the Ndb as a time-related database.

Database maintenance is an ongoing effort by a staff of NCC nutritionists trained in database maintenance procedures. Ongoing maintenance involve

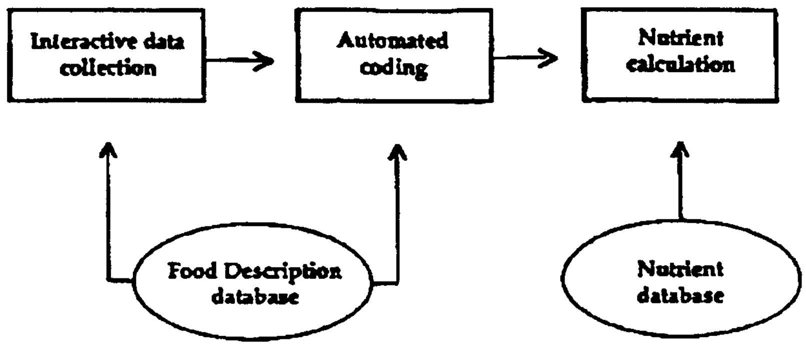



Figure 1. Major steps and databases required for automated collection and processing of dietary data using the Minnesota Nutrition Data System (NDS)

incorporation of new data from the USDA Human Nutrition Information Service and information on commercial products from manufacturers, as well as information from the literature, from foreign food composition tables and from recipe books. A comprehensive listing of the sources of data for the NCC Ndb has been reported by Schakel et al (8). A new version of the Ndb is released every 6 to 12 months; 16 versions of the database have been released over the past decade.

To ensure that intake data collected in the past can be recalculated on any previous or future version of the Ndb, the following guidelines are strictly adhered to by the database nutritionists:

1. All new foods and new formulations are added to all previous versions. This allows studies using a "frozen" database to keep current with the marketplace.
2. No foods are ever deleted from the Ndb. Products no longer on the market continue to be updated when the database is expanded to include additional nutrients or other non-nutrient food components.
3. Modifications to nutrient values are added only to the current version of the database. Even though better analytic data become available, any modifications to a database already in use by a study would confound the interpretation of study results.
4. For products that are reformulated, multiple date-related nutrient values are maintained in the database. The computer accesses the appropriate nutrients based on the date the food was consumed.
5. When an entry in the database is expanded into two or more entries, the entry that is the most representative of past intakes retains the original code number.

Adherence to these guidelines ensures that dietary data collected at any point in time can be recalculated to take advantage of all updates to the Ndb. These guidelines have been in place at the NCC since 1983, and many US and Canadian studies have taken advantage of the ability to compare nutrient data over time by recalculating previously coded data using an updated version of the Ndb.

**Description and maintenance of the NCC Food Description database (FDdb).**

The FDdb maintained at the NCC has been described in detail elsewhere (5,6). Briefly, the FDdb consists of two components. One component is a hierarchical structure of food descriptions that drives the interactive food identification process. This hierarchy of food descriptions is based on the way ordinary people describe foods. A series of menu-type screens prompt the user for as much detail as is necessary to permit the computer to automatically assign the appropriate code numbers and gram amounts. The level of food description becomes progressively more detailed as the hierarchy is traversed for a given food item. These detailed descriptions may include brand names as well as methods of food preparation.

The second component of the FDdb is the information required to automatically assign code numbers and gram weights to the foods identified through the interactive prompting described above. Such information includes all of the common synonyms for describing a particular food, the recipes or formulations for describing combination foods, designation of ingredients that require further description (eg, the type of fat used to prepare a food), and the defaults to be assigned if the necessary level of detail cannot be provided by the subject. In addition, the FDdb must be able to provide adequate information for converting amounts as described by the subject into gram weights. Amount conversions may require densities, raw to cooked yields, edible portion factors, geometric shape conversions, or weights of food specific units.

The more traditional non-automated procedures for processing dietary data result in substantial loss of descriptive detail at the coding level. For example, most brand name products are assigned code numbers corresponding to generically described foods in the Ndb. Thus, brand identification is lost at the time of coding, since only the generic code numbers and amounts are entered into the computer. If a product had been assigned an incorrect code number, the only way to correct the error would be to manually recode the data. Because manual coding is very labor intensive, recoding is not a feasible option unless the coding process is automated. Automated coding is possible only if the computer is able to capture all of the food description detail required for coding. Thus, the FDdb described above incorporates the features needed for automated recoding at a later time. However, automated recoding is possible only if successive versions of the FDdb are compatible. (Procedures to ensure compatibility between successive versions of the FDdb are discussed in the next section.)

All FDdb modifications are made to the current working version of the database. Updates include additions of new products that enter the market, modifications for reformulated products, and changes reflecting new or improved data. Foods no longer on the market are deleted only after a reasonable period that takes into account the potential time lag between purchase and consumption of the product.

To maintain currency with the marketplace and changing food consumption patterns, ongoing studies are encouraged to collect data using only the most recent version of the FDdb. However, all previous versions used by the study must be retained to permit editing of food intake records collected in the past, since each record is tied to that version of the FDdb on which it was originally collected.

**Development and maintenance of a Multi-Version Food Description database.**

To enhance the comparability of dietary data collected over time, the NCC has developed a Multi-Version Food Description database (MVdb). Use of the MVdb eliminates the logistical problems of maintaining many different versions of the FDdb for editing food records collected in the past. All existing versions of the FDdb are collapsed into a single database, thus eliminating the large amount of redundancy among the separate databases. The MVdb automatically accesses the appropriate version each time a dietary intake record is edited. Most importantly, the MVdb permits coding changes to be retroactive to previous versions of the FDdb. Thus, food intake records collected in the past are able to take advantage of subsequent corrections and additions to the FDdb.

Maintenance of the MVdb requires that every update to the database is assigned one of three _troactivity codes.y One code is used to indicate that the change is retroactive to all previous versions of the FDdb; a second code is used to indicate that the change is not retroactive to any previous versions; and a third code, representing a specific version of the FDdb, is used to indicate that the change is retroactive back to the designated version. NCC database nutritionists use the following guidelines for selection of retroactivity codes:

1. Improved numeric data are always retroactive to all versions of the FDdb. For example, a more accurate value for a food density or for a raw to cooked conversion factor needs to be reflected in the recalculated data whenever food intake records collected with a previous version of the database are recalculated on an updated version.
2. Changes in the food supply are retroactive to the version reflecting the time the change occurred. For example, the addition of a new commercial product is retroactive to that version of the database most closely matching the date the product was introduced into the marketplace. This allows the editing of past food intake records for which the new product was reported and temporarily assigned a missing food status because the product had not yet been added to the FDdb.
3. Deletions are never retroactive. Even though a product is no longer on the market, the product must remain in all previous databases to permit editing of food intake records collected in the past for which the product may have been reported. However,

deletions are permitted in a newer version of the database reflecting the foods available in the current food supply.

Adherence to these guidelines ensures currency for data collection and comparability of coding and nutrient calculation for dietary data collected at any point in time.

## Summary

Time-related databases that permit comparison of food and nutrient intakes over time are essential for nutrition monitoring, as well as for long-term studies investigating diet-health relationships. It is possible to maintain nutrient and coding databases in a manner that will permit routine updating to reflect the changing marketplace and the availability of improved data while simultaneously providing data analysis stability for long-term studies and trend analysis of food consumption data.

Comparison of food and nutrient intake data over a long time period is possible only if certain requirements are met. These requirements include the following: 1) the database used for food identification must be frequently updated to maintain currency with the dynamic marketplace and changing food preparation and food consumption patterns; 2) detailed food descriptions, including any relevant information on brand names or food preparation methods, must be captured; 3) coding procedures must be automated so that the raw data (ie, the detailed food descriptions) collected at different times can be recoded using the same coding procedures; and 4) all nutrient intakes must be calculated using the same nutrient database. Compliance with these requirements will ensure comparability of food intake data collected at any point in time, permitting trend analysis of food and nutrient intakes with minimal confounding due to differences in coding procedures or nutrient databases.

## References


1. Healthy People 2000, National Health Promotion and Disease Prevention Objectives. DHHS Publication No. (PHS) 91-50212, Supt of Documents, US Government Printing Office, Washington DC, 1991.

2. Sievert, YA, Schakel, SF, and Buzzard, IM: Development and maintenance of a nutrient database for clinical trials. Controlled Clinical Trials, 10:416-425, 1989.

3. Dennis B, Ernst N, Hjortland M, Tillotson J, Grambsch V: The NHLBI Nutrition Data System. J Am Dietet Assoc 77:641-647, 1980.

4. Tillotson JL, Gorder DD, Kassim N: Nutrition data collection in the Multiple Risk Factor Intervention Trial (MRFIT). J Am Dietet Assoc 78:235-240. 1981.

5. Feskanich D, Buzzard IM, Welch BT, Asp EH, Dieleman LS, Chong KR Bartsch GE: Comparison of a computerized and a manual method of food coding for nutrient intake studies. J Am Dietet Assoc 88:1263-7, 1988.

6. Feskanich,D, Sielaff BH, Chong K, and Buzzard IM: Computerized collection and analysis of dietary intake information. Computer Methods and Programs in Biomedicine 30:47-57 1989.

7. McDowell M, Briefel RR, Warren RA, Buzzard IM, Feskanich D, Gardner SH: The Dietary Data Collection System: an automated interview and coding system for NHANES III. In: Proceedings of the 14th National Nutrient Databank Conference, June 19-21, 1989, Iowa City, IA; Stumbo PJ, ed; The CBORD Group, Inc, Ithaca, NY; 1990. pp125-131.

8. Schakel SF, Sievert YA, and Buzzard IM: Sources of data for developing and maintaining a nutrient database. J Am Dietet Assoc 88:1268-71, 1988.

# Obtaining and Using Industry Data

**Sandra J. Bartholmey, Ph.D.**
**Gerber Products Company**
**Fremont, Michigan**

## INTRODUCTION

As a representative of a baby food company for whom accurate nutrient data are a high priority, I would like to give you an idea about why we maintain a nutrient data bank and how we generate nutrient data. These comments may help answer some of the questions we receive about the accuracy of our nutrient data and why complete nutrient data on products are not always available.

For the record, all of the nutrient analyses of Gerber baby foods have been done in-house since we started nutrient analyses of our products in the 1940's. Our nutrient data bank goes back to 1973, when labeling regulations were first implemented. I want to give credit to Wesley Meeuwsen for helping me with this presentation. Wes is the person that answers most of our questions about about our nutrient data bank.

## WHY MAINTAIN A DATA BANK?

The primary purpose of our nutrient data bank is to provide the best possible data for nutrition labeling of our products. Since baby foods are considered Foods for Special Dietary Use by FDA, they are subject to mandatory nutrition labeling regulations. Compliance with government regulations on labeling is important to our business, so we maintain our data bank as the basis for our label declarations for both FDA and USDA regulated products.

## SAMPLING AND ANALYSIS

Gerber analytical laboratories analyze each product every year for all nutrients declared on the label. We follow the FDA protocol for sampling of products, that is, a sample is defined as a composite of twelve containers pulled at random throughout a one-day pack period. Each composite, therefore, represents a single "sample". We consider data from three "samples" the minimum on which to base a label declaration.

In addition, each product is analyzed every three years for nutrients **not** declared on the label, i.e., some of the B vitamins and trace minerals.

For cereal products that contain added nutrients, such as iron, analyses are run every four months on a composite of all production packs made during that four-month sampling period.

All of these pieces of data go into the nutrient data bank and are used for several purposes.

## HOW IS THE DATA BANK USED?

The primary purpose of maintaining the nutrient data bank is for mandatory nutrition labeling. Any other uses of the data bank must be subordinate to this first priority. Programs are in place to convert the raw data into amounts per serving and percent U.S.RDA for determining label declarations. Nutrient data are reviewed on a regularly scheduled basis and compared to label data. Adjustments of label declarations are made as needed.

The nutrient data bank is also used to annually update **Nutrient Values** handbook, one of Gerber's most widely used publications by health professionals and consumers. **Nutrient Values** contains more nutrient data than is found on the labels; and the data are presented in units per 100 grams as well as in units per serving.

Nutrient data on nutrients that do not appear on the label or in **Nutrient Values**, some of the vitamins and trace minerals, are maintained on a less rigorous schedule than the label nutrients. We use this information to answer special requests from health professionals working out special diets for patients with special needs, from researchers using Gerber foods in a study who want to calculate nutrient intake data, and from data bank builders who want to include baby foods in their data sets.

It should be made clear that nutrients appearing on the food label are the primary focus. We do not monitor the other nutrients as rigorously because they are not needed for label declarations. Analyses are costly to run, and it is difficult to justify maintaining an expensive analytical program to management if there is no compelling reason to do so, particularly in the current climate of cost reduction and down-sizing. Requests for specialized data are justified for legal or public relations reasons and to aid Product Development in product design. Lack of time and business impact are the primary reasons for the low priority given to sharing nutrient data.

A final use for our nutrient data bank is to support the Gerber Infant Nutrition Survey (INS) which we conduct every three or four years. The INS provides us with dietary intake data for babies two to eighteen months of age. The intake data for Gerber products can be converted into nutrient intake data based on information residing in our nutrient data bank. Nutrient intake data on other food items in the survey data bank are collected based on information from secondary sources rather than direct analysis.

The INS can help us follow trends in infant feeding practices and answer questions about what babies eat, when they eat it, and how much they eat. Outside of Gerber, this information is shared with nutritionists, pediatric audiences and government agencies that have a use for food intake data on infants. Internally, it is interpreted by nutritionists for Marketing and Product Development.

## IMPACT OF NEW LABELING REGULATIONS

The FDA labeling proposals and the Nutrition Labeling and Educations Act of 1990 have already significantly increased the analytical work load in our laboratories in preparation for the May, 1993, implementation date. Whether foods for infants and toddlers will be required to label all the proposed nutrients is under review. Even if the exemption sought for baby and toddler foods from labeling of select nutrients is gained, we will continue to collect the data to publish in a context that allows for educating the consumer in the different dietary needs of infants and toddlers compared to adults.

## SUMMARY

The major use of the Gerber nutrient data bank is to maintain compliance with government labeling regulations. This conference offers us, the users, managers and builders of data bases, the opportunity to understand the similarities as well as the differences in our priorities and goals in the use and maintenance of nutrient data bases.

agers
nder-
n our
ce of

# Obtaining and Using Industry Data

**Jenny Chin**
**Best Foods, CPC International**
**Union, New Jersey**

**Best Foods Nutrition Policy**

Best Foods recognizes its obligation to provide nutrition information about its food products to the consumer and health professional. Best Foods has provided nutrition information on its products since 1973 when nutrition labeling became voluntary. Skippy peanut butter and Karo syrup were the first products to voluntarily bear such label information. Since then it has been Best Foods policy to include basic nutritional data on all products unless the size of the package makes it physically impossible, for example our Golden Griddle syrups, or if the product is not a significant source of nutrients; for example our Argo and Kingsfords corn starch. Nevertheless, detailed information for all Best Foods products is available upon request.

**Generation of Data**

Nutrient data for our products are generated in several ways. The data may be calculated from the product formula and ingredient information or the product and ingredients may be chemically analyzed for key nutrients.

For new products, where there are limited data available, it may be necessary to calculate a theoretical nutrient profile using the formula and ingredient information. Ingredient information may be obtained from a variety of sources: such as standard food composition references like the USDA's Handbook 8 and its revised editions, Bowes and Church's **Food Values of Portions Commonly Used** and McCance and Widdowson's **Composition of Foods.** Other sources of ingredient information are obtained by analyzing individual ingredients in our analytical testing laboratories or using outside contract labs to handle analyses that we don't routinely perform. Ingredient information is also obtained from the vendors of the individual ingredients. The information supplied by a vendor is usually submitted in the form of specification sheets. Although the information is useful, many times it is incomplete making it necessary to contact the vendor for any missing data. This can be extremely time consuming and frustrating. The moisture, cholesterol or sodium contents of an ingredient are examples of nutritional components that are rarely provided.

Each time the nutrient profile of a new ingredient is compiled, the data are entered and stored in our computerized nutrient database system: The Best Foods Nutrient Data Base System. Once the nutrient composition of all ingredients in the product is in the database, the nutrient profile of the product can be calculated. The method used by our computerized nutrient database system is the summing method. For instance, if the sodium content of each ingredient in a product is known as well as the percent contribution that each ingredient provides to the product, by summing or adding the sodium contents of each ingredient, the total amount of sodium in the product can be determined. Our nutrient data base system can calculate and report up to 72 different nutritional components in a product or recipe.

To substantiate the theoretical values and to develop a database for a new product, data are generated by chemically analyzing samples which are produced in the laboratory, at the pilot plant or production plant level. At least 6 individual samples are analyzed for proximate composition, that is moisture, protein, fat and ash; the carbohydrate content is determined by difference. Sodium, cholesterol, vitamins and minerals are also analyzed if present at a significant level. The analytical procedures used are those of the Association of Official Analytical Chemists (AOAC) or equivalent methods. The analytical data generated are reviewed by our Statistics Department to determine the mean and standard deviation of each nutrient as well as the values which would be representative of the product. The natural variation in ingredients, the difference in precision between analytical methods, as well as variation which may be introduced by different production plants is taken into consideration.

Once a product is established and manufactured on a regular basis, a sampling plan is developed as part of the Best Foods Nutrient Assay Control Program. The Best Foods Nutrient Assay Control Program was developed to ensure that all values that are declared on our product labels are in compliance with FDA guidelines. Each product has its own sampling plan and is analyzed for key nutrients on a regular basis. For example, each month, four samples of each type of Skippy peanut butter are collected from every production plant. The four samples are composited and then analyzed for protein, moisture, total fat, fatty acid composition, niacin and sodium. Pasta products are analyzed for proximate composition as well as vitamins and minerals. Our dressing products are analyzed for total fat, fatty acid composition, sodium and cholesterol. Well over six thousand different analyses are run on our products each year to monitor compliance with the nutrition labels. These analyses are in addition to any tests run by our Quality Control Department.

**Use of Data**

The data generated for our products are used in a variety of ways. The primary uses are: 1) to develop nutrition information for our product labels, 2) ensure compliance with label values, 3) develop our product nutritional data sheets, 4) develop menus and recipes showing how to use our products, and 5) provide data to the United States Department of Agriculture's National Nutrient Data Bank.

First, the data that are generated for our product labels must be declared following rounding and format rules outlined in the Food and Drug Administration's Code of Federal Regulations section 101.9. titled: The nutrition labeling of foods.

The calorie content of a product is declared to the nearest 2 Calorie increment up to 20 Calories, to the nearest 5 Calorie increment from 20 to 50 Calories, and to the nearest 10 Calorie increment if there are more than 50 Calories per serving. The protein, carbohydrate and fat contents are declared to the nearest gram; Cholesterol is declared to the nearest 5 mg increment if there is more than 2 milligrams per serving; if there is less than 2, then 0 is declared.

Sodium is declared to the nearest 5 mg increment from 5 up to 140 milligrams per serving, and to the nearest 10 mg increment above 140 milligrams per serving; if there is less than 5 milligrams, 0 is declared. Second, the data generated in our Nutrient Assay Control Program are used to ensure compliance with label values. Third, data are also used to prepare and update our product nutritional data sheets which are comprised of the Nutrition Information Per Serving (Table 1) and the Approximate Composition (Table 2) data sheets. The Nutrition Information Per Serving data sheet contains the information just described, that is, values that are rounded according to FDA's regulations. This is the information found on our product labels. The Approximate Composition data sheet expresses the absolute (i.e. not rounded) nutrient values per 100 grams of the product, as well as the amount in common household units of measure. Data for components not listed on the nutrition panel such as moisture and ash content are listed on the Approximate Composition data sheet.

Fourth, the product nutrient data are used in the development of recipes for consumer and health professional materials. These are two examples of recipe booklets where the nutrient profile of each recipe is provided. And finally, data generated by our Nutrient Assay Control Program for our dressings, oils, margarines and peanut butter products are sent periodically to the United States Department of Agriculture's National Nutrient Data Bank.

This is just a brief overview of our program describing how nutrient composition data for our food products are generated, and how that information is used. We recognize that it is a valuable tool used to assess dietary intakes in the health care and academic areas. We will continue to generate and provide this information and welcome any suggestions you may have on how we can continue to meet your needs.

TABLE 1

**NUTRITION LABELING REPORT: MAZOLA MARGARINE**
**Nutrition Information Per Serving**

| | |
|---|---|
| Serving size: 1 tablespoon (14 grams) | |
| Servings per container (pound): 32 | |
| Calories | 100 |
| Carbohydrate | 0 grams |
| Fat | 11 grams |
| Percent of calories from fat* | 99% |
| Polyunsaturated* | 4 grams |
| Saturated* | 2 grams |
| Cholesterol* | 0 milligrams (0 mg/100 g) |
| Sodium | 100 milligrams |

Percentage of U.S. Recommended Daily Allowances (U.S. RDA)

| | |
|---|---|
| Vitamin A | 10 |
| Vitamin D | 15 |

Contains less than 2 percent of the U.S. RDA of protein, vitamin C, thiamine, riboflavin, niacin, calcium, iron.

TABLE 2
**APPROXIMATE COMPOSITION REPORT: MAZOLA MARGARINE**

| Component | In 100 Grams | In Tablespoon | In 1 Cup |
|---|---|---|---|
| Weight (grams) | 100.0 | 14.0 | 229.0 |
| Moisture (grams) | 15.9 | 2.2 | 36.4 |
| Protein (N x 6.25) (grams) | 0.2 | None | 0.5 |
| Fat, total (grams) | 80.7 | 11.3 | 184.8 |
| Triglycerides (grams) | 79.5 | 11.1 | 182.1 |
| Polyunsaturates, total (grams) | 28.6 | 4.0 | 65.5 |
| Cis, Cis only (grams)* | 26.7 | 3.7 | 61.1 |
| Monounsaturates (grams) | 37.2 | 5.2 | 85.2 |
| Saturates, total (grams)* | 13.7 | 1.9 | 31.4 |
| C12:0 to C18:0 (grams)* | 13.7 | 1.9 | 31.4 |
| P/S value | 2.1 | 2.1 | 2.1 |
| Unsaponifiable matter (grams) | 1.2 | 0.2 | 2.7 |
| Cholesterol (milligrams) | None | None | None |
| Phytosterols (milligrams) | 500 | 70 | 1145 |
| Tocopherols (milligrams) | 56 | 8 | 128 |
| Carbohydrate, available (grams) | 1.3 | 0.2 | 2.9 |
| Ash (grams) | 1.9 | 0.3 | 4.4 |
| Sodium (milligrams) | 710 | 100 | 1625 |
| Vitamin A (International Units) | 3570 | 500 | 8180 |
| Vitamin D (International Units) | 430 | 60 | 980 |
| Energy (Calories) | | 720 | 1001650 |

# Obtaining and Using Industry Data

**Frances H. Seligson, Ph.D.**
**Hershey Foods Corporation**
**Hershey, Pennsylvania**

## INTRODUCTION

The current practices at Hershey Foods for obtaining nutrient information for its products and sharing this information with the public -- either through the food label or in brochures -- were covered in an earlier presentation. The impact of NLEA (The Nutrition Labeling and Education Act) and FDA (Food and Drug Administration) proposed regulations on these practices was also discussed. It is inevitable that the builders of nutrient databases will feel the impact of these labeling initiatives in their future requests for data.

This presentation provides more of the details and issues involved in compiling and sharing nutrient information about our products and, when relevant, the impact of mandatory nutrition labeling (Table 1).

TABLE 1

**Compiling and Sharing Nutrient Data**

- Data Generation
  - Sample size
  - Sample collection
  - Sample analysis
- Data Reporting
  - Unit of measure
  - Decimal places
  - Rounding
- Data Sharing
  - Company brochures
  - Client forms
- Issues
  - Sample size
  - Serving size
  - Brand names and trademarks
  - Competing priorities

## DATA GENERATION

The nutrition information which appears on the food label ideally is derived from the analytical results of seven production lots (each lot is represented by a composite of 12 samples). Information for nutrients which are provided voluntarily in brochures is usually based on the analytical results from at least three production lot composites.

The collection of a production lot composite requires an interruption in the daily routine and scheduling of plant operations. In the case of multiple manufacturing sites, a decision must be made about how to sample. The nutrition labeling for new products in test market often is based on pilot plant production runs. The production lot composites are prepared by Hershey Foods and sent to a contract laboratory for analysis of all nutrients, except for fatty acids which are determined internally.

The increased expense and product sampling load created by mandatory nutrition labeling will probably result in the analysis of fewer production lot composites, fewer optional nutrients, and fatty acids by a contract laboratory.

## DATA REPORTING

The analytical data for each production lot composite is entered onto a spread sheet in metric units per 100 grams (except vitamin A is entered in IU) to 2-3 decimal places as reported by the lab. A program calculates an analytical mean and 2 standard deviations about that mean. The mean and 2 standard deviations are used to develop nutrient information for the food label and for our nutrient information brochures.

Nutrient information on the food label and in our brochure for consumers is expressed per serving size; the values reflect standard deviation corrections to the analytical mean and rounding according to current regulations. The nutrient information in our brochure for health professionals is the analytical mean per 100 grams. Values in the latter brochure are reported to the nearest whole number or 1-2 decimal places depending on the nutrient; energy is rounded to the nearest 10 calories. Fatty acid data currently is provided in separate fact sheets but will be incorporated into future brochures.

## DATA SHARING

Hershey Foods frequently receives requests for nutrient information from nutrient database builders and the staff tries to accommodate these requests. Some want information expressed as the mean per 100 grams; some want it per serving; and others want it both ways. We usually fulfill these requests by providing both of our nutrition information brochures, sometimes supplemented with our fact sheets for cholesterol and fatty acid content.

Often we receive preprinted product forms from database builders with a request to "fill in the blanks" or "check the numbers for accuracy." This does not pose a problem if only a few products are involved; but it is a burden when there are inches of forms. We recently received a request which consisted of 250 pages of product forms, and some of the products were not even ours!

## ISSUES

In providing nutrient information to database builders, certain issues seem to resurface on a regular basis. One is sample size. We are sometimes asked to share the number of sample analyses used to arrive at a nutrient value. As a matter of practicality and confidentiality, we do not provide that information.

Another issue is serving size. Some request uncorrected nutrient information per serving (i.e., not converted and rounded as shown on the food label). It is difficult to respond to these requests because (1) the data are not readily available in that format and (2) the serving size for a given brand varies as a function of bar weight under current regulations. For example, our standard milk chocolate bar is available in seven sizes, from 0.27 oz. for miniatures to 7.0 oz. for giant; serving size varies from 1.0 to 1.75 oz. (Table 2). FDA's proposal for standard serving sizes will address this particular issue but will then create its own set of problems.

A serious issue is protection of our brand names and registered trademarks. It is absolutely critical that nutrient database builders print our brand names and registered trademarks as exactly indicated. In fact, they are obligated to do so.

A last and chronic issue is that of competing priorities. Most of Corporate America is staffed leanly today. And we are stretched to the limit, especially now, in tracking the various labeling issues, preparing for mandatory nutrition labeling, and handling ongoing internal requests for services. Fulfilling requests for nutrient information from the builders of databases, unfortunately, is getting buried under an ever growing stack of immediate business priorities.

TABLE 2

**Bar and Serving Sizes for HERSHEY'S Milk Chocolate Bar**

| *Bar Size* | *Bar Weight* (oz) | *Serving Size* (oz) | *Servings Per Bar* |
|---|---|---|---|
| Miniatures | 0.27 | N/A | N/A |
| Snack | 0.50 | 1.0 | 1/2 |
| Standard | 1.55 | 1.55 | 1 |
| Fund Raiser | 2.2 | 1.1 | 2 |
| Big Block | 2.6 | 1.3 | 2 |
| King | 4.0 | 1.0 | 4 |
| Giant | 7.0 | 1.75 | 4 |

# Obtaining and Using Industry Data

**Mary Stevens**
**University of Minnesota**
**Minneapolis, Minnesota**

**Rationale for Collection of Brand Specific Product Information**

In 1974, in collaboration with the National Heart Lung and Blood Institute, the Nutrition Coordinating Center (NCC) developed a highly standardized system for collecting and analyzing dietary data (1). This system was initially developed for medical research studies investigating the relationships between diet and cardiovascular diseases. Dietary interviewers for the studies were trained to ask subjects to identify fat containing foods, such as margarine, oil, shortening, cookies and crackers, by brand so that the resulting nutrient calculations would reflect the total fat, fatty acids, and cholesterol content of foods eaten.

We have continued to collect nutrient data for brand name products for the past 17 years to accommodate increasing levels of nutrient specificity required by the studies using the NCC system. We have expanded the system to include specificity for other nutrients of research interest, such as sodium, dietary fiber, sucrose, and a number of vitamins and minerals.

**Obtaining Industry Data**

Currently, NCC collects brand specific nutrient and serving size information from the food industry for over 6,000 products. To keep pace with a marketplace in which more than 12,000 brand name and fast food products are introduced annually, two nutritionists devote the major portion of their time to obtaining and processing these data. Approximately 350-400 letters are sent to manufacturers each year.

We have a mailing list of manufacturers that has been compiled over the years and which we continually update with new information as needed. The Thomas Food Industry Register (2) is used to identify the names and addresses of manufacturers of many new products appearing on dietary intake records sent to us by the studies using our services.

An update log has been developed to determine when each manufacturer should be contacted. This log contains information including the specific products of interest, the date of the last update, and the date of the next scheduled update, as well as any special notations which may be made by the nutritionist.

Computer-generated letters and NCC Nutrient Data Forms are sent annually to request updated product information. If the manufacturer does not respond to our letter, we follow up with a second letter or a telephone call to try to obtain the information. If necessary, we will check grocery stores for label information or ask clients in other states to send us information unavailable locally. Canadian clients using the NCC system have been especially helpful in sending us updated information about the Canadian products in our database. On occasion, we purchase products to obtain serving size information that may not be on the label, such as the weight or dimensions of a candy bar or cookie.

**Processing of Brand Name Product Data**

Until recently, we have stored data received from manufacturers in a paper file. This file contains data for more than 6,000 products. Over the past year we have developed a computerized brand name database in which we can now store data as we receive them from manufacturers. A detailed description of the components of the Brand Name Database is provided elsewhere in these Proceedings (3). Upon receipt of information from a manufacturer, relevant data are entered into the Brand Name Database, and the information is used to update our Nutrient and Food Description Databases.

Databases for Shelf Labeling:

# "GTI Data Document Systems"

**Karen Falk**
**Data Document Systems**
**Industrial Airport, Kansas**

**Introduction**

Graphic Technologies Inc.- Data Document Systems is the manufacturer of bar coded labels. In 1986, we were asked by our grocery customers, to provide nutrition information on the labels that we already provided for their daily operations. We first attended this conference in 1986 while researching their request for nutrition labeling. When we returned to the Mid West, we contacted the University of Kansas Medical Center for data assistance. Nutrient data there was used for research and for average daily and weekly intake records. Each item in the database contained many fields of nutrient data.

For the production of a shelf-edge label, the final product is completely different and the data management is also different. This morning I will tell you something of the issues involved in the production of a shelf-edge label that displays nutrition information, such as;

1. Data Management Issues
2. The Media of the Shelf-edge Label
3. The Involvement of the Food and Drug Administration
4. The Value to the Consumer
5. The Value to the Nutrition Community.

**Data Management Issues**

For the production of a *product* label or *shelf-edge* label, fewer nutrient categories are of issue because of the limited space, but additional data categories become necessary for the management of the large amount of product-specific data. Fields like the 10-digit UPC number become critical to channeling the data to its' desired destination. The UPC code or Universal Product Code, identifies who made the product and specifically which product by kind and size. To that number, can be attached any number of facts, like product movement (by the manufacturer and the grocer), the price (by the grocer), and any number of demographic considerations when coordinated with a grocer's customer file, such as "Who buys cat food". For the purposes of our discussion this morning, we will only consider one kind of UPC-related data - nutrient data.

**The Media of the Shelf-edge Label**

The media of the shelf-edge label provides a very effective method or communicating. Here is an example of a basic shelf-edge label. You are probably aware of the existence of these labels in your grocery stores.

The original purpose of this label was to help the grocer communicate pricing information to the customer, unit pricing (required in some states) and to help the grocer reorder quickly and accurately by providing a scannable bar code. This bar code is read instantly by a light projection scanner and translated to a number: the UPC or Universal Product Code number.

We are all experienced in a certain kind of poverty - a poverty of time. "Fast" becomes a necessity of many of our activities, especially routine ones like grocery shopping. So when you want to communicate something in the grocery store you must do it *clearly, briefly and attractively,* or your message will not take.

**The FDA is Involved**

When nutrition information is the issue, the shelf-edge label is viewed as an extension of the product label and regulated like the product label by the Food and Drug Administration. The concern of the FDA is that the consumer be presented with consistent terminology. For the manufacturers of a *product* or the manufacturers of shelf-edge labels to use descriptors such as Low Sodium, Low Calorie, and Source of Fiber, FDA-established criteria must be met;

"Low Calorie", 40 calories or less per serving and no more than .4 calories per gram

"Low Sodium", 140 mg. or less per serving

"Low Fat", 2 gm. or less per serving, and no more than 10% fat on a dry weight basis

"Low Cholesterol", 20 mg. or less per serving

"Source of Calcium", 100 mg. or more per serving

"Source of Fiber", 2 gm. or more per serving.

At present, serving size is as stated by the manufacturer. You may be aware that new regulations for nutrition labeling are being formulated and will become part of the Code of Federal Regulations in November of this year. The issue of "serving size" may become regulated within product categories. It is also possible that the "Low Fat" and "Low Cholesterol" categories will be combined, meaning that before a product can be called "Low Cholesterol", it must fit both the Low Cholesterol and Low Fat criteria.

Source of Calcium is based on the USRDA for Calcium and USRDA's are changing to "Daily Values", so there may be a slight change in this category.

The primary focus of the new legislation is the product label, but since the shelf-edge label is seen as an extension of the product label, it will touch this area also.

**Value to the Consumer**

The value of displaying nutrition information on a shelf edge label is that you have the opportunity to help the consumer make educated choices about the foods and beverages that they consume, *at the point-of-purchase,* where according to the Point Of Purchase Advertising Institute, 66% of all buying decisions are made.

The data directly reaches the consumer; not as it relates to a daily or weekly average, but as it relates to a serving of the product considered for purchase.

When a product label says that the product contains 120 milligrams of sodium, some consumers need help interpreting that amount into its' relevance to their dietary concerns. Therefore, when 120 milligrams can be described as "Low Sodium", the sodium level in this product becomes *relevant* to the consumer.

Modification is a concept that is critical to making good dietary choices. It is, however, a concept that is difficult to effectively communicate with the space allowed on a shelf-edge label. The shelf-edge label is the first step, providing a very elementary education but the only education some consumers will choose to receive. The second is an informational brochure which helps fill in criteria and category information. The third and final step, and one we encourage our grocers to consider, is the involvement of local dieticians in store tours or group information sessions. The consumer then has an opportunity to ask questions of and become acquainted with the services available through the local professional.

**Value to the Hispanic Consumer**

We are especially pleased to be making this information available in Spanish. At an FDA hearing in San Antonio, we heard Dietician Rosario Hamilton ask for Spanish nutrition information from food manufacturers. San Antonio has an Hispanic population of 70%. Many of the people she counsels speak English only as a second language.

The 1990 Census reported that 9% of the population of the United States is of Hispanic origin. In the Hispanic household, to a greater degree than the Anglo household, the women are the decision makers about food purchases, preparation and family health. They are also less likely to be educated than the men of the house, and are therefore more likely to be communicating primarily in Spanish. When Spanish is the first language, product labels in English are of little use.

Manufacturers may have a difficult time responding to this need because of all that is already required on the product label. Here a shelf-edge label becomes the communications media that crosses that language barrier with some very important information.

**Value to the Nutrition Community**

According to the latest study repeated annually by the Food Marketing Institute, *96% of grocery shoppers consider nutrition an important consideration when making a food or beverage purchase.* This is the time hoped for, when the consumer makes food choices based on nutrition value, rather than taste, trends and the seduction of clever advertising.

Certainly, changes are ahead, but it is a dynamic and exciting time for nutrition education. We are pleased to provide an education tool for our grocers and for the nutrition community.

Using the Menu Census Survey to Estimate Dietary Intake:

# Post Market Surveillance of Aspartame

**I. Jack Abrams**
**Market Research Corp. of America**
**Northbrook, Illinois**

## INTRODUCTION

MRCA has been monitoring the actual consumption of Aspartame by individuals from their diets since its approval in 1981. This task was mandated by The Food & Drug Administration (FDA) as a condition of its approval of Aspartame as a food additive, as published in the Federal Register on July 24, 1981. One objective of this monitoring was to compare the actual use with the maximum projected consumption level of 34 milligrams per kilogram of body weight of the eater, per day, at the 99th percentile for the total sample. This estimate was computed by MRCA in 1975, using food consumption data from the Menu Census Survey of 1972-73, and the concentrations of Aspartame that were proposed by General Foods Corporation in packets for table-top use, in various powdered mixes, and in several liquid products including carbonated soft drinks. It was included in the petition which was submitted to FDA by General Foods Corporation in March 1976. Thus, MRCA's work in estimating the intake of Aspartame began about seventeen years ago.

The methodology for using food consumption data from Menu Census Surveys to estimate the intake of substances from the diet was originally developed, in the early 1970's, by the National Academy of Sciences GRAS Review Committee Phase I. It has been refined substantially since then, in continued work with The Food & Drug Administration, with the National Academy of Sciences GRAS Review Committee Phase II and Phase III, and with many commercial organizations. Over these years, the Menu Census Surveys were used to estimate the frequency distributions of the intake of close to 2,000 substances from the diet. In order to understand this methodology, it is necessary to describe in detail the Menu Census Surveys.

## THE MENU CENSUS SURVEYS

The Menu Census Surveys provide an in-depth continuing record of food and beverage preparation and consumption by U.S. households and individuals. Although the primary use of these data is for marketing and product development, a growing use has developed in the past 20 years for estimating the intake of direct and indirect food additives, by both The Food and Drug Administration and by commercial organizations, in order to meet FDA's regulatory and review requirements. In addition, these data have been used by commercial organizations to estimate the intake of vitamins and minerals from the diet, in order to support nutritional claims, or to explore new product opportunities. Some of their key applications, and typical user groups, are shown in the following table:

| APPLICATIONS | USERS |
|---|---|
| **National Food Trends** | *Top Management*<br>*Mergers & Acquisitions* |
| **Defining the Marketplace:**<br>Market Size Estimates<br>Competitive Framework<br>Product Positioning<br>Copy Theme Development | *Product/Marketing Management*<br>*Advertising Management & Agencies*<br>*Legal Department* |
| **Managing Brands:**<br>Line Extension<br>Recipe Ideas<br>Tie-In Promotions<br>Ad Themes | *Product/Marketing Management*<br>*Advertising Management & Agencies*<br>*Nutritionists/Home Economists* |
| **Evaluating New Product Opportunities:**<br>Define Markets & Segments<br>Estimate Potential<br>Target Audiences | *New Product Development* |

| | |
|---|---|
| **Forecasting:** | *Strategic Planning Groups* |
| Long & Short Range | *Marketing Management* |
| Detect Trends - Early On | *Pronuct Management* |
| **Intake Studies:** | *Food Development & Technology* |
| Nutrition | *Regulatory Affairs* |
| Food Additives | *Government Agencies* |
| | *Nutritionists* |

MRCA has been tracking all food preparation and consumption at-home and away-from-home through the Menu Census Surveys since 1957. These surveys were conducted once in five years from 1957 to 1977, and then continuously from 1980 forward. All foods are reported, except the use of table salt, pepper, and tap water. The surveys are currently based on nationally representative rotating samples of 500 households per quarter, or 2,000 households per year, containing about 5,100 members. Each household reports all food preparation and consumption, daily, by mail, in 14 consecutive daily diaries. The households are distributed uniformly throughout the year, with about five or six new households starting their two-week reporting period each day of the year. All diaries are completed by the homemakers, who are also long-term members of MRCA's National Consumer Panel (NCP), the Weekly Purchase Diary Panel, who are therefore experienced in reporting their purchases of food products, in great detail, via diaries by mail.

## THE DAILY DIARY

As shown in ATTACHMENT I, pages 1 to 5, each daily diary provides the following information:

1. A detail description of each dish eaten, and items added to it at the table
2. At-home or away-from-home
3. At breakfast, lunch, or dinner meals; or at morning, afternoon, evening, or bedtime snack eating occasions
4. The position of the dish in the meal
5. Which household members ate the dish, and each item added to it at the table

For all dishes eaten at-home, information is provided on:

1. The number of guests who ate it, by children vs. adults
2. First time vs. leftover serving
3. Method of preparation and appliance used
4. Brand names of commercially prepared products
5. Form as obtained, such as fresh, frozen, canned, etc.
6. Packaging material in contact with the food
7. For homemade dishes, the reporting includes a detail description of every product used as an ingredient, fats and oils used as agents for frying, or flour for dusting breadboards. For each ingredient, the diary provides the brand name, form as obtained, packaging material, and whether the ingredient was itself a leftover.
8. Who ate the meal together, at what time, and if it was a special occasion.

For foods eaten away-from-home, information is provided on:

1. The type of place, such as at friends, school, restaurant, lunch counter, etc.
2. The name of the food service facility
3. If from a vending machine
4. If eaten at the place where it was obtained.

## DIETS, PSYCHOGRAPHICS, AND HOUSEHOLD DEMOGRAPHICS

A separate questionnaire, administered following the 14th day of reporting, provides, for each household member, detail information on age, sex, pregnancy status, weight, height, special diets followed, reason for the diet, foods encouraged or discouraged from eating, use of table salt, and the consumption of vitamin and mineral supplements, including kind, potency, amount and frequency, as is shown in ATTACHMENT II, pages 1 and 2.

Another questionnaire, completed only by the homemaker, covers attitudes, awareness, and interests in a wide range of subjects dealing with lifestyle, food preparation, cooking skills, nutrition, food additives or preservatives, low-calorie products, sugar substitutes, and the use of information printed on the labels, as is shown in ATTACHMENT III, pages 1 and 2.

In addition, an extensive set of demographic classifications is available for each household, obtained as part of their participation in NCP, the Weekly Purchase Diary Panel.

## INTAKE STUDIES

An intake study for a given substance usually includes Frequency Distribution Reports of the Daily Intake of the Substance by Individuals in several age groups, in milligrams (MG) and in milligrams per kilogram of body weight (MPK) of the eaters. Also provided are the corresponding Sources of Mean Intake Reports, which show the contribution of each specific food to the total overall mean intake of the substance, by the same age groups.

The intake study for a given substance is based on a detailed listing of all the foods which contain the substance, or which the manufacturer plans to include in the petition to FDA for the use of this substance.

This list is prepared using the detail Menu Census Food Classification Code Book. At the same time, the manufacturer also provides the actual or the proposed concentrations of the substance in each food item on the list. The amount of the substance consumed by eating any food item on the list is computed by multiplying the concentration of the substance by the quantity of the food eaten.

## QUANTITIES OF FOOD EATEN

Note that the 14 daily diaries provide only the incidence of eating each food product by an individual, but not the quantity eaten by each person, since reporting quantities for 14 days would be too burdensome to the homemaker. Instead, the average grams per eating occasion have been calculated from the most recently available USDA National Food Consumption Survey (NFCS), for persons grouped by age and sex, using a linear "smoothing" procedure on these estimates when needed.

Since the USDA Survey provides grams of foods for end dishes as-eaten, these estimated average amounts per eating occasion are used only to quantify dishes as-eaten, such as milk when consumed as a beverage, or sugar when added to coffee or tea, or oil used as a salad dressing. When the products are used at-home as ingredients or as frying agents in preparing other foods, the amounts consumed are computed as percentages of the corresponding amounts of the end dishes in which they were used. These percentages are based on estimates obtained from standard recipes.

## INTAKE AMOUNTS

The quantity of the substance consumed by a given individual, from all eating occasions of a given item of food on the list, in a given day, is thus calculated by multiplying the number of such eatings by the average grams per eating occasion, and by the concentration of the substance in that food. This calculation is displayed below, in the case of Aspartame:

The Estimated Total Intake of Aspartame, by a Given Person, on a Given Day, from All Eatings of a Given APM Containing Food Item Eaten that Day, by that Person, is the Result of the Multiplication of the Following Three Terms:

| | | | | |
|---|---|---|---|---|
| Number of *times* an APM containing item of food was eaten on that day by that person | * | Average number of *grams* per eating occasion of that food for a person of that age & sex group | * | Number of *milligrams* of APM per gram of that food as eaten |

These estimated intakes of the substance from different food items are then aggregated, for all eatings of all food items by that same person, **separately for each day,** to provide estimates of intake on a **Person-Day Basis,** treating each day for each person as an independent observation. Intakes are also accumulated for each person **over all 14 days combined,** to provide estimates on a **14-Day Average Daily Basis.**

## INTAKE STUDY REPORTS

Frequency distributions of the average daily intake of the substance are then tabulated by age groups, in five percentile increments, for the eaters only, and separately for the total sample of eaters plus non-eaters, from the lowest to the highest, with finer breakouts for the heavier eaters of the substance above the 95th percentile, usually showing the 97.5th, the 99th, the 99.5th, and the 100th percentiles. Separate reports show these intakes in milligrams (MG), vs in milligrams per kilogram (MPK) of the body weight of the eater. Reports are usually produced for the intake of the substance from specific food sub-categories, vs. from broader food categories, and in total from all foods combined.

Multiple reports are frequently produced, separately on a **Person-Day Basis** vs. on a **14-Day Average-Daily Basis,** as needed to differentiate analyses of the potential **"acute"** effects from the **"chronic"** effects of the consumption of the substance.

Corresponding reports are provided for the Sources of Mean Intake, which break out the contributions to the mean intake, by age groups, of each specific food item included in the original list of concentrations of the substance.

## CONCENTRATIONS BY BRAND AND FLAVOR

In addition, since the Menu Census Surveys identify the brands of each commercial product, reports can be produced tracking the "actual intake" of a food additive, once it is approved and introduced into the market, and as it expands its distribution over time; instead of the "prospective intake" of the substance, which is estimated when a petition to FDA is first prepared for the use of the substance in specified foods. This is the case, for example, with the Post Marketing Surveillance - Phase II of Aspartame, in which MRCA has been tracking the intake of APM since April 1984, using its actual concentrations in each brand, as reported by their manufacturers.

## NUTRITION STUDIES

The intake of any nutrient in the food is estimated using essentially a similar procedure, with the concentrations coming from a standard nutrient composition database. For this purpose, each food in the Menu Census Survey is assigned its corresponding code in the nutrient composition database, and the associated amount of the nutrient per 100 grams of the food is treated as the appropriate "concentration".

The reports for nutrition studies usually show the contributions of each food category and sub-category to the total daily intake of each vitamin or mineral, for individuals classified by age and sex groups. Additional reports distribute the intakes of nutrients by meal occasion, or by various food consumption patterns of the individuals in the study. These intakes are usually reported by age and sex groups, in absolute quantities, or as averages of the percentages of the Recommended Daily Allowances (RDA's) for the corresponding nutrients, or by ranges of the percent RDA's, and the like.

## SPECIAL DIETS

Since the Menu Census Study contains information on any special diets which a person may follow, reports are also produced comparing the intake of food additives, or of vitamins and minerals, by persons who are, for example, on a diet to reduce weight, or to control diabetes.

## HOMEMAKER'S ATTITUDES

Using the responses of the homemakers to the attitudes, awareness, and interests questionnaire, which has been administered to all Menu Census households since 1972, MRCA has been able to classify these homemakers by their concern about low-calories, nutrition, food additives, prepared foods, etc., and then correlate these classifications with the frequency of consumption of various food categories by individuals in the same households. Using the same classifications, it is possible to trend the nutrient intakes by the homemakers, and by other household members, over the past 19 years, a period during which interest in proper nutrition has grown substantially. Similar trends can be produced for selected food additives whose consumption may be correlated with specific homemaker's attitudes, such as towards fat or cholesterol, sugar, low-calorie sugar substitutes, caffeine, preservatives, and the like.

## THE POST MARKETING SURVEILLANCE OF ASPARTAME - AN INTAKE STUDY EXAMPLE

As stated earlier, the Post Marketing Surveillance (PMS) of Aspartame, PHASE I, was begun in January 1982. It was designed to track the percentage of children 0 to 12 years old who consumed any food containing added Aspartame during an average 14-day period. PHASE I was to continue until the level of about 30% was reached, corresponding to the previously existing level of exposure by children to Saccharin containing products. PHASE II would then begin, in which the frequency distributions of the "actual" amount of Aspartame consumed through the diet will be tracked on a quarterly basis.

## PHASE I - EXPOSURE BY CHILDREN 0-12 YEARS OLD

Chart A displays the growth in the exposure of children to Aspartame containing products from April 1982 through June 1984. For the first year, less than 2.1% of children 0-12 years old consumed any products containing Aspartame in an average two-week period. Beginning with April 1983, the exposure to Aspartame grew rapidly, and reached a level of about 24% by June 1984. In fact, the percentage of children 2-5 years old consuming added Aspartame reached 30% by June 1984, thus initiating the Post Market Survey PHASE II.

## PHASE II - FREQUENCY DISTRIBUTIONS OF THE INTAKE OF ASPARTAME

PHASE II of the Post Marketing Surveillance of Aspartame began with the second quarter of 1984. Quarterly and annual frequency distribution reports were provided through 1987, and only annual reports were continued since then. They were provided in milligrams and in milligrams per kilogram of body weight of the eater, on a Person-Day basis, and on a 14-day Average Daily basis (14-Day Avg). Reports were produced by age and sex groups, as well as for persons on a Diet to Reduce Weight (Reducers) and on a Diabetic Diet (Diabetics).

The attached charts show quarterly trends in the 90th percentiles of intakes of APM, for eaters only, in MPK, for children 2-5 vs. 6-12 years old, and for reducers vs. diabetics. Each trend line is compared to that for the Total Sample. They start with the second quarter of 1984, and go to the fourth quarter of 1987, when quarterly reports were stopped, and only annual reports were continued. CHARTS B through E are on a person-day basis, and CHARTS F through I on a 14-day average daily basis.

### Person-Day Intakes

As can be seen in CHART B, the intake by 2-5 year olds is substantially higher than by the total sample, fluctuating from a low of 8.5 MPK in the fourth quarter of 1984 to a peak of 19 MPK in the first quarter of 1986, and declining to about 10 MPK in the third and fourth quarters of 1987. During this same period, the average intake by the total sample was below one-half that of children 2-5 years old, and hovered around 5 MPK. This difference is characteristic of intakes expressed in MPK, since small children eat more food in relation to their body weight than adults, and therefore also most food additives.

As can be expected, the intake by children 6-12 year olds, shown in CHART C, is somewhat lower than that by 2-5 years old. It rose from a low of 6.4 MPK in the third quarter of 1984, to about 9 MPK in 1985, with a single extreme value of 17 MPK in the fourth quarter of 1985, and then settled down to about 7 MPK in 1986 and 1987.

The person-day intake of individuals on a diet to reduce weight, and also of those on a diabetic diet, display an unexpected pattern, since they are both essentially lower, instead of higher, than the intake by the total sample, as shown in CHARTS D and E. However, this pattern is once again due to the unit of measure, namely MPK's, since almost all reducers and diabetics are adults, who weight more, and who eat less per kilogram of body weight, than children. Thus, the resulting average MPK intake for the total sample is increased by the higher MPK intakes of the "lighter" children with relatively greater food intake.

The high fluctuations in the APM intake by diabetics is caused primarily by the small sample size for this group.

### 14-Day Average Daily Intake

CHARTS F through I show a strikingly different pattern in the intake of APM on a 14-day average daily basis than did CHARTS B through E on a person-day basis. Unlike the person-day trends, those for the 14-day average daily intake by 2-5 and 6-12 year olds dropped sharply in the fourth quarter of 1986, and then fluctuated around the intake by the total sample for the next five quarters. Obviously, this difference must be caused by the fact that 14-day averages reflect any existing patterns of consistency or regularity in the intake of APM, whereas such factors are not incorporated into the calculations of the short-term person-day intakes. Specifically, only one-third to one-half as many children eat any APM containing foods as adults do on a day-by-day basis, compared to over eight-tenths to nine-tenths as many children as adults do on a 14-day basis. Thus, more children are infrequent eaters of APM containing foods, and therefore eat less APM on a "long-term" average daily basis as compared to adults.

Similarly, a significant change is reflected in the pattern of APM intake for persons on a reducing or on a diabetic diet, as is shown in CHARTS H and I. Contrary to their person-day intakes, which were essentially below those for the total sample, their 14-day average daily intakes are now larger than those for the total sample, and they also reveal an upward trend in these intakes from 1984 through 1987. These trends may reflect the tendency by these two groups to adopt a pattern of regular consumption of APM containing food products, which is not shared by children, and possibly neither by other adults.

### CONCLUSION

The Menu Census Survey has proven over the years to be a very effective instrument for estimating the intake of food additives from the diet for both "chronic" as well as "acute" levels of exposure. The richness of the database has supported the extreme demands of estimating intake from foods which are "ready-to-eat", as well as from those which are used by the homemaker only as ingredients or cooking agents for preparing other dishes. The extended 14-day period of observation is indispensable in estimating the long-term average intake of nutrients from the diet; the "chronic" exposure to direct food additives; to food animal drug residues; as well as to contaminants in the food, such as naturally occurring lead, or that which migrates into the food from sodered cans.

Furthermore, once a substance has been approved by FDA, the same Menu Census database has frequently been used to set priorities for the introduction of the product into different food categories, to define the segments of the prospective market, to estimate trends, to explore for new uses in additional food items, and for various other market research and marketing purposes, as well as for Post Marketing Surveillance of newly approved food additives.

**CHART A: APM EATERS IN A 14-DAY PERIOD**

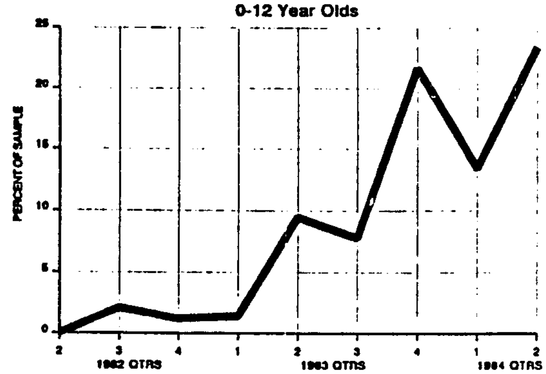


**CHART B: PERSON-DAY BY 2-5 YEAR OLDS**
90th Pctl of Daily APM Intake by Eaters Only

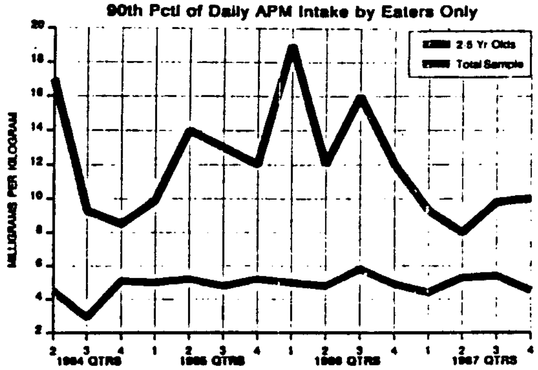



**CHART C: PERSON-DAY BY 6-12 YEAR OLDS**
90th Pctl of Daily APM Intake by Eaters Only

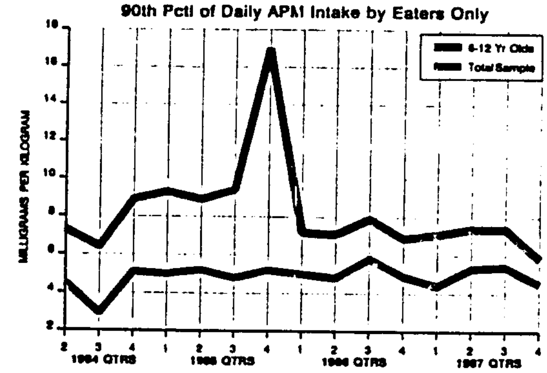



**CHART D: PERSON-DAY BY REDUCERS**
90th Pctl of Daily APM Intake by Eaters Only

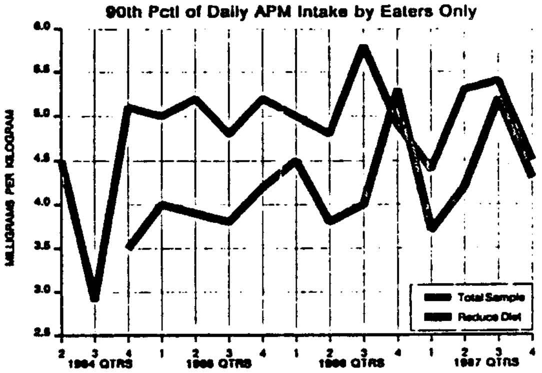



**CHART E: PERSON-DAY BY DIABETICS**
90th Pctl of Daily APM Intake by Eaters Only

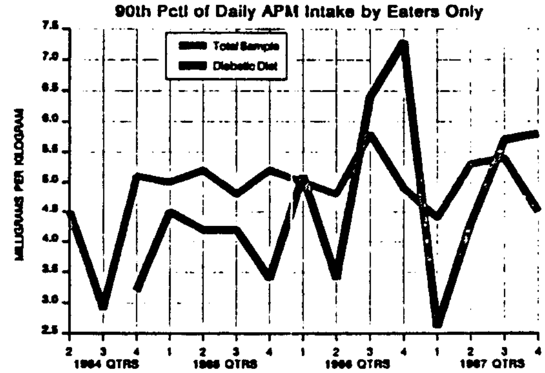



**CHART F: 14-DAY AVERAGE BY 2-5 YR. OLDS**
90th Pctl of Daily APM Intake by Eaters Only

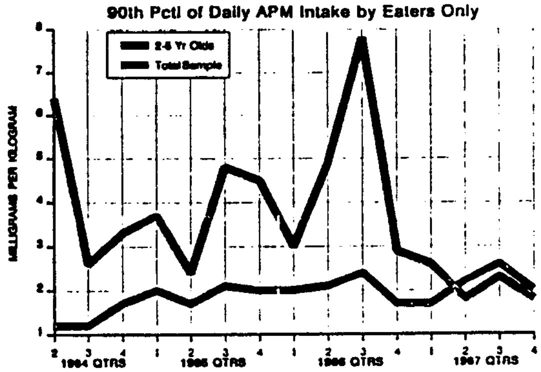



**CHART G: 14-DAY AVERAGE BY 6-12 YR OLDS**
90th Pctl of Daily APM Intake by Eaters Only

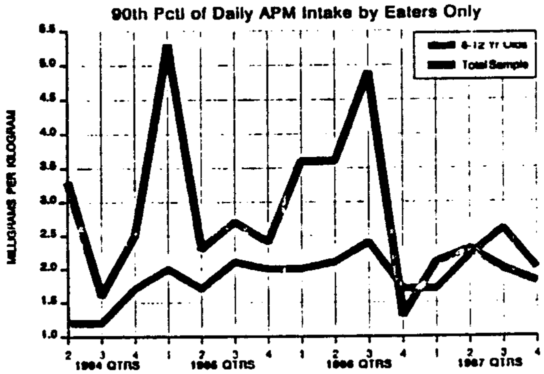



**CHART H: 14-DAY AVERAGE BY REDUCERS**
90th Pctl of Daily APM Intake by Eaters Only

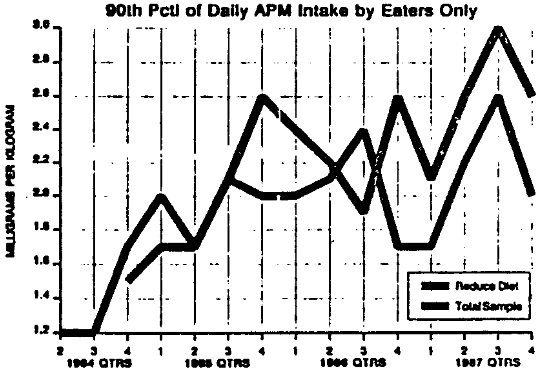
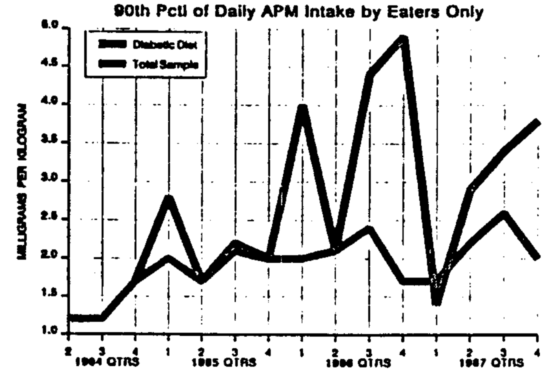
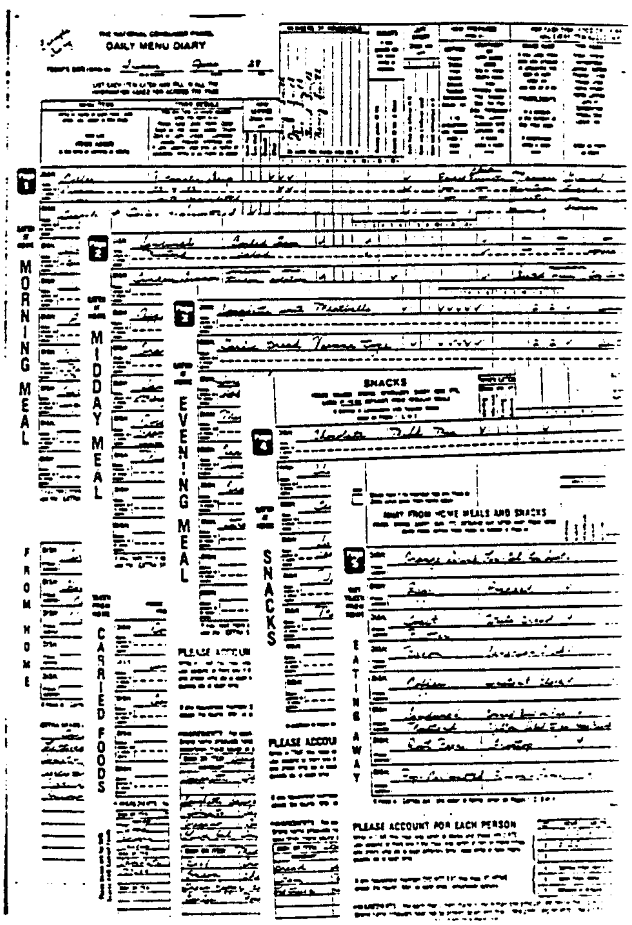
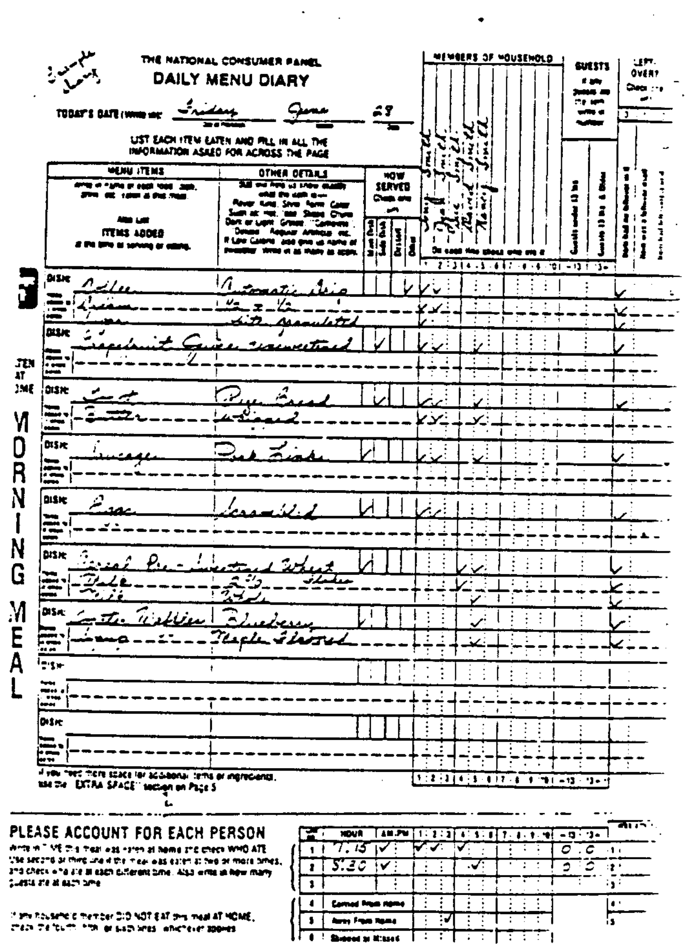
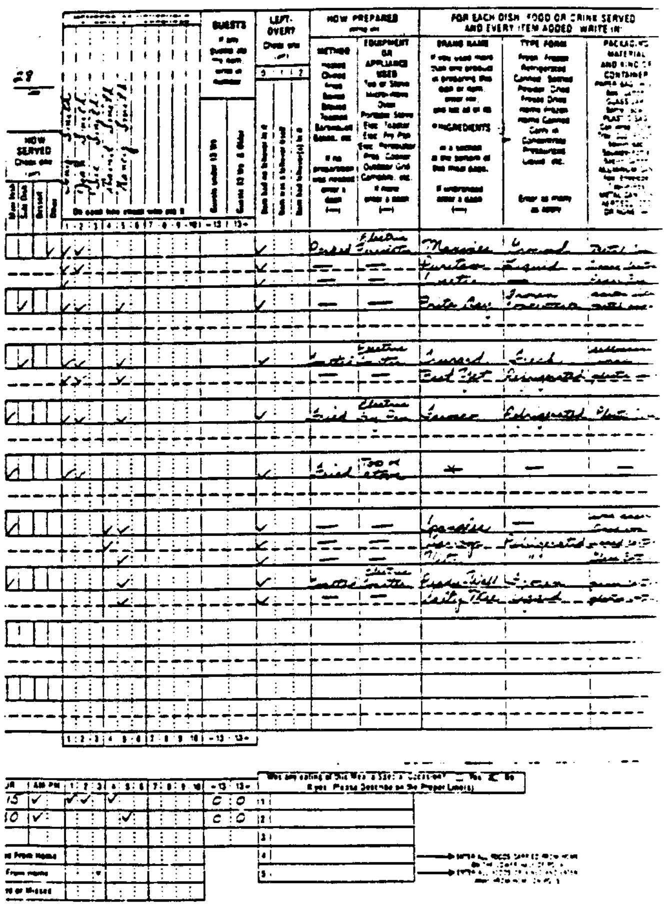


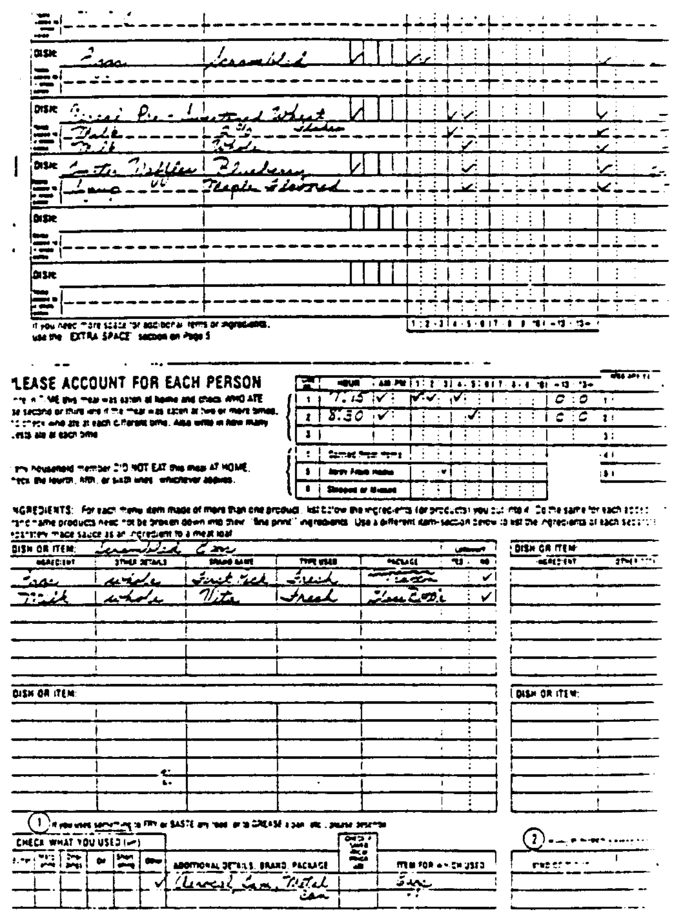

**CHART I: 14-DAY AVERAGE BY DIABETICS**
90th Pctl of Daily APM Intake by Eaters Only

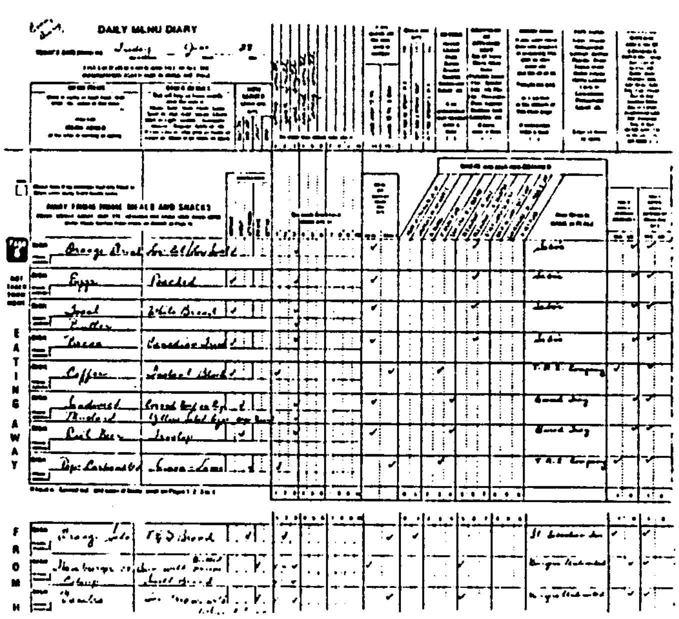





THE NATIONAL CONSUMER PANEL
DAILY MENU DIARY
TODAY'S DATE
LIST EACH ITEM EATEN AND FILL IN ALL THE INFORMATION ASKED FOR ACROSS THE PAGE
MENU ITEMS
ITEMS ADDED
OTHER DETAILS
HOW SERVED
MEMBERS OF HOUSEHOLD
GUESTS
MORNING MEAL
PLEASE ACCOUNT FOR EACH PERSON


HOW SERVED
GUESTS
LEFT-OVER?
HOW PREPARED
FOR EACH DISH, FOOD OR DRINK SERVED AND EVERY ITEM ADDED, WRITE IN:


PLEASE ACCOUNT FOR EACH PERSON
DISH OR ITEM:
CHECK WHAT YOU USED


DAILY MENU DIARY

PANEL NUMBER

PLEASE ANSWER ALL QUESTIONS FOR EACH MEMBER OF YOUR HOUSEHOLD

| | Please write in Month & Year of Birth for each person listed And Add names of persons not listed | [illegible] | On Diet? | If you answered YES to the previous question, please check below the appropriate boxes which best describe the diet for this person, and write in details as required: Name of Diet / Reason / Rx? / Food Avoided or Encouraged | Table Salt | Vitamin or Mineral Supplements |
|---|---|---|---|---|---|---|
| 1 | MONTH YEAR | | | | | |
| 2 | MONTH YEAR | | | | | |
| 3 | MONTH YEAR | | | | | |
| 4 | MONTH YEAR | | | | | |
| 5 | MONTH YEAR | | | | | |
| 6 | MONTH YEAR | | | | | |
| 7 | MONTH YEAR | | | | | |
| 8 | MONTH YEAR | | | | | |
| 9 | MONTH YEAR | | | | | |
| 10 | MONTH YEAR | | | | | |

IF YES ENTER DETAILS ON BACK PAGE

PLEASE COMPLETE THE SECTION BELOW TO TELL US MORE ABOUT THE VITAMINS AND MINERALS YOU AND YOUR HOUSEHOLD MEMBERS TAKE:

| | Complete Brand Name and Manufacturer | KIND | [illegible] | FORM | Potency | Name of Users | Amount Used | How Often Does Each Person Use |
|---|---|---|---|---|---|---|---|---|
| 1ST PRODUCT | | | | | | | | |
| 2ND PRODUCT | | | | | | | | |
| 3RD PRODUCT | | | | | | | | |
| 4TH PRODUCT | | | | | | | | |
| 5TH PRODUCT | | | | | | | | |

Please fold and return in the enclosed stamped envelope

THE NATIONAL CONSUMER PANEL

## Special Quiz About Your Household Members

CMC 89/90

THIS QUIZ IS WORTH 50 POINTS

Dear Panel Member,

We have been very pleased with your Daily Menu Diaries. They have been arriving regularly and we are now putting together all the information about your meals (along with the same records from thousands of other families) to prepare reports on the foods families in this country serve and eat. The study will be a success, thanks to your fine help.

Some additional information is still very necessary so that we can better understand the uses of food.

This is all covered in the SPECIAL QUIZ printed inside this folder. Yes, some of the questions may seem personal but we sincerely hope you will answer them all. As you know, all of this information will be treated confidentially, as is all of your reporting.

The Quiz is mainly concerned with special foods eaten because of diets, etc. We hope you will find it interesting and that you will take a few moments to complete it. Note the section about Vitamins and Minerals on the back page. After you've completed the Quiz, mail it back to us in the enclosed stamped envelope. Please do this as soon as possible.

Thanks again for your fine reporting. We are glad we can count on you.

Cordially,

Doris Reed

DR:ch

PS The points you will have earned for keeping the 14 Daily Menu Diaries plus the 50 for the Quiz will be shown on your Diary labels within the next month or two.



ATTACHMENT II-2

| | Disagree Completely | Disagree Somewhat | Disagree Slightly | Neither Agree Nor Disagree | Agree Slightly | Agree Somewhat | Agree Completely |
|---|---|---|---|---|---|---|---|
| 89 I really enjoy cooking | ☐ | ☐ | ☐ | ☐ | ☐ | ☐ | ☐ |
| 90 I enjoy spending time preparing a meal with fancy dishes | ☐ | ☐ | ☐ | ☐ | ☐ | ☐ | ☐ |
| 91. Children are sure to dislike nutritious foods if they are told to eat them by their parents | ☐ | ☐ | ☐ | ☐ | ☐ | ☐ | ☐ |
| 92 Members of my family often refuse to eat their vegetables | ☐ | ☐ | ☐ | ☐ | ☐ | ☐ | ☐ |
| 93 My family is not easy to please | ☐ | ☐ | ☐ | ☐ | ☐ | ☐ | ☐ |
| 94 I frequently worry about whether the meals I serve are really nourishing | ☐ | ☐ | ☐ | ☐ | ☐ | ☐ | ☐ |
| 95 I feel that if I give my family a large variety of foods, they will get the proper nutrition | ☐ | ☐ | ☐ | ☐ | ☐ | ☐ | ☐ |
| 96 I make every possible effort to see that my family eats really nourishing foods | ☐ | ☐ | ☐ | ☐ | ☐ | ☐ | ☐ |
| 97. I know more about nutrition than most people do | ☐ | ☐ | ☐ | ☐ | ☐ | ☐ | ☐ |
| 98 Red meat is better for your health than fish | ☐ | ☐ | ☐ | ☐ | ☐ | ☐ | ☐ |
| 99 Chicken is as good for you as fish | ☐ | ☐ | ☐ | ☐ | ☐ | ☐ | ☐ |
| 100 I always see to it that my family takes vitamins | ☐ | ☐ | ☐ | ☐ | ☐ | ☐ | ☐ |
| 101 Our family sometimes discusses various foods and their food values so that they understand nutrition better | ☐ | ☐ | ☐ | ☐ | ☐ | ☐ | ☐ |
| 102 I prepare each meal to be nutritionally balanced | ☐ | ☐ | ☐ | ☐ | ☐ | ☐ | ☐ |
| 103 I buy fish because it is very nutritious | ☐ | ☐ | ☐ | ☐ | ☐ | ☐ | ☐ |
| 104 I keep my recipes in an organized file | ☐ | ☐ | ☐ | ☐ | ☐ | ☐ | ☐ |
| 105 I'd like to save more recipes but find it hard to keep track of them | ☐ | ☐ | ☐ | ☐ | ☐ | ☐ | ☐ |
| 106 I'm much more willing to try a new recipe when someone I know tried it out and liked it | ☐ | ☐ | ☐ | ☐ | ☐ | ☐ | ☐ |
| 107 I get ideas about new dishes from eating in restaurants | ☐ | ☐ | ☐ | ☐ | ☐ | ☐ | ☐ |
| 108 I often exchange recipes with friends and relatives | ☐ | ☐ | ☐ | ☐ | ☐ | ☐ | ☐ |
| 109 I frequently collect recipes from magazines | ☐ | ☐ | ☐ | ☐ | ☐ | ☐ | ☐ |



| | Disagree Completely | Disagree Somewhat | Disagree Slightly | Neither Agree Nor Disagree | Agree Slightly | Agree Somewhat | Agree Completely |
|---|---|---|---|---|---|---|---|
| 110 I frequently collect recipes from the food section of the newspaper | ☐ | ☐ | ☐ | ☐ | ☐ | ☐ | ☐ |
| 111 I frequently send away for new recipes | ☐ | ☐ | ☐ | ☐ | ☐ | ☐ | ☐ |
| 112 I prepare foreign dishes more often than I used to | ☐ | ☐ | ☐ | ☐ | ☐ | ☐ | ☐ |
| 113 Foreign dishes are just the thing to serve when you have important company | ☐ | ☐ | ☐ | ☐ | ☐ | ☐ | ☐ |
| 114 Sometimes I buy food or other products imported from foreign countries even when they cost a little more | ☐ | ☐ | ☐ | ☐ | ☐ | ☐ | ☐ |
| 115 I collect recipes for low calorie foods | ☐ | ☐ | ☐ | ☐ | ☐ | ☐ | ☐ |
| 116 Most foods are as nutritious when they are served as leftovers | ☐ | ☐ | ☐ | ☐ | ☐ | ☐ | ☐ |
| 117 Most foods are as tasty when they are served as leftovers | ☐ | ☐ | ☐ | ☐ | ☐ | ☐ | ☐ |
| 118 Though I try to make use of leftovers, I throw most of them away | ☐ | ☐ | ☐ | ☐ | ☐ | ☐ | ☐ |
| 119 My family can't tell the difference between food that is freshly prepared and that which is left over from a previous meal | ☐ | ☐ | ☐ | ☐ | ☐ | ☐ | ☐ |
| 120 In our house, leftovers are often eaten as snacks | ☐ | ☐ | ☐ | ☐ | ☐ | ☐ | ☐ |
| 121. Meats baked in the oven taste better | ☐ | ☐ | ☐ | ☐ | ☐ | ☐ | ☐ |
| 122 I use my oven frequently | ☐ | ☐ | ☐ | ☐ | ☐ | ☐ | ☐ |
| 123 My family loves fried foods | ☐ | ☐ | ☐ | ☐ | ☐ | ☐ | ☐ |
| 124 Health foods are not worth their extra cost | ☐ | ☐ | ☐ | ☐ | ☐ | ☐ | ☐ |
| 125 The term "natural" means only those foods grown without use of artificial chemicals or fertilizers | ☐ | ☐ | ☐ | ☐ | ☐ | ☐ | ☐ |
| 126 So-called health foods are not any better for you than the well known brands | ☐ | ☐ | ☐ | ☐ | ☐ | ☐ | ☐ |
| 127 I would like to know more about what various foods contain and how to avoid possible harmful ingredients | ☐ | ☐ | ☐ | ☐ | ☐ | ☐ | ☐ |
| 128 You don't really have to worry about nutrition if you include plenty of fresh fruit in your diet | ☐ | ☐ | ☐ | ☐ | ☐ | ☐ | ☐ |





**SECTION F** The following questions concern your overall attitudes and opinions toward low-calorie sugar substitutes, both those which are available for your own use, and those which come as ingredients in foods and beverages. For each statement, please check (√) the appropriate box which indicates how you feel about these products in general.

| | Disagree Completely | Disagree Somewhat | Disagree Slightly | Neither Agree Nor Disagree | Agree Slightly | Agree Somewhat | Agree Completely |
|---|---|---|---|---|---|---|---|
| 163 Low-calorie sugar substitutes assist in good nutrition | ☐ | ☐ | ☐ | ☐ | ☐ | ☐ | ☐ |
| 164 Low-calorie sugar substitutes should be used only by those people who cannot eat sugar because of health or medical reasons | ☐ | ☐ | ☐ | ☐ | ☐ | ☐ | ☐ |
| 165 You do not get cavities from eating low-calorie sugar substitutes | ☐ | ☐ | ☐ | ☐ | ☐ | ☐ | ☐ |
| 166 I wish there were more products containing low-calorie sugar substitutes instead of sugar | ☐ | ☐ | ☐ | ☐ | ☐ | ☐ | ☐ |
| 167 It is all right for children to eat low-calorie sugar substitutes | ☐ | ☐ | ☐ | ☐ | ☐ | ☐ | ☐ |
| 168 Low-calorie sugar substitutes do not have any long-term harmful health effects | ☐ | ☐ | ☐ | ☐ | ☐ | ☐ | ☐ |
| 169 Most low-calorie sugar substitutes taste good | ☐ | ☐ | ☐ | ☐ | ☐ | ☐ | ☐ |
| 170 Some low-calorie sugar substitutes have an unpleasant aftertaste | ☐ | ☐ | ☐ | ☐ | ☐ | ☐ | ☐ |
| 171 I am concerned that some low-calorie sugar substitutes can cause brain damage | ☐ | ☐ | ☐ | ☐ | ☐ | ☐ | ☐ |
| 172 You can add low-calorie sugar substitutes to any food or beverage, either hot or cold | ☐ | ☐ | ☐ | ☐ | ☐ | ☐ | ☐ |
| 173 Whenever possible, I prefer using low-calorie sugar substitutes rather than sugar in recipes or at the table | ☐ | ☐ | ☐ | ☐ | ☐ | ☐ | ☐ |
| 174 It is all right for pregnant women to eat low-calorie sugar substitutes | ☐ | ☐ | ☐ | ☐ | ☐ | ☐ | ☐ |
| 175 You can use most low-calorie sugar substitutes for cooking and baking | ☐ | ☐ | ☐ | ☐ | ☐ | ☐ | ☐ |
| 176 I am concerned that low-calorie sugar substitutes can cause cancer | ☐ | ☐ | ☐ | ☐ | ☐ | ☐ | ☐ |



| | Disagree Completely | Disagree Somewhat | Disagree Slightly | Neither Agree Nor Disagree | Agree Slightly | Agree Somewhat | Agree Completely |
|---|---|---|---|---|---|---|---|
| 177 I am willing to pay more for products containing low calorie sugar substitutes in place of sugar | ☐ | ☐ | ☐ | ☐ | ☐ | ☐ | ☐ |
| 178 It is all right for older people to eat low-calorie sugar substitutes | ☐ | ☐ | ☐ | ☐ | ☐ | ☐ | ☐ |
| 179 All low calorie sugar substitutes are artificial | ☐ | ☐ | ☐ | ☐ | ☐ | ☐ | ☐ |

**SECTION G** The following questions are about your use of certain information shown on food labels.

| | Never (1) | Rarely (2) | Sometimes (3) | Often (4) | Always (5) |
|---|---|---|---|---|---|
| 180 When shopping for food, to what extent do you use information shown on the label to decide what brand of a product to buy? Check (√) one | ☐ | ☐ | ☐ | ☐ | ☐ |
| 181. In your efforts to plan a balanced dietary intake for your family, to what extent do you use the nutrition information available on the labels of some food products? Check (√) one. | ☐ | ☐ | ☐ | ☐ | ☐ |
| 182. When buying canned goods, to what extent do you buy brands according to the label's solid weight declaration? (Solid weight is the weight of the contents exclusive of liquid.) Check (√) one. | ☐ | ☐ | ☐ | ☐ | ☐ |

**SECTION H** For each of the purposes listed below, please check (√) how often you use the ingredient statement or the nutritional information panel available on food labels to:

| | Never (1) | Rarely (2) | Sometimes (3) | Often (4) | Always (5) |
|---|---|---|---|---|---|
| 183 Avoid or reduce sugar | ☐ | ☐ | ☐ | ☐ | ☐ |
| 184. Select products with no sodium or low sodium | ☐ | ☐ | ☐ | ☐ | ☐ |
| 185. Avoid products that use chemical preservatives, artificial colors or additives | ☐ | ☐ | ☐ | ☐ | ☐ |
| 186 Avoid or reduce amount of fat | ☐ | ☐ | ☐ | ☐ | ☐ |
| 187 Control calories | ☐ | ☐ | ☐ | ☐ | ☐ |
| 188 Increase protein | ☐ | ☐ | ☐ | ☐ | ☐ |
| 189 Increase vitamins/minerals | ☐ | ☐ | ☐ | ☐ | ☐ |
| 190 Avoid cholesterol | ☐ | ☐ | ☐ | ☐ | ☐ |

# Langual for Database Users

**Elizabeth C. Smith, M.S.L.S.**
**Food and Drug Adminstration**
**Washington, District of Columbia**

Finding nutritional, consumption or toxicological information on food can be as difficult as finding one's way through a rat maze. The Center for Food Safety and Applied Nutrition (CFSAN) of the U.S. Food and Drug Administration (FDA) has developed an excellent information retrieval system that links dissimilar databases through a common descriptive language. The information retrieval system is known as the **Food Monitoring Database** and the descriptive system is called **Langual**. The name comes from French words meaning "Universal Language for Food." Langual has been presented at previous National Nutrient Databank Conferences. For a detailed description of its construction and features, see page 93 of the Proceedings of the Fourteenth National Nutrient Databank Conference, June 19-21, 1989, University of Iowa, Iowa City, Iowa.

Langual is a standardized vocabulary for food product description. It is composed of fourteen different viewpoints or factors such as "Product Type" or "Extent of Heat Treatment." The factors contains numerous descriptive terms, each of which may be used to retrieve information about a food product. Langual provides assistance to users in a number of ways. Definitions explain what a term is or how it is meant to be used. Synonyms for scientific nomenclature or vernacular usage are available; French synonyms were added in the past year. Retrieval terms are arranged in a hierarchy, which arrays terms conceptually from broader to narrower.

Langual is used for information retrieval in CFSAN's Food Monitoring Database (FMDB). It is stored at the Parklawn Computer Center in Rockville, Maryland, on an IBM 3090 K64, using Model 204 Database Management System. Nine diverse food databases have been indexed by using Langual. Six of these files are from sources outside the FDA. They are the USDA Nutrient Database for Standardized Reference (Handbook 8); the 1987/88 Nationwide Food Consumption Survey; food names from the **Codex Alimentarius**; a carotenoid food file; a French food file and a Greek food file. The three remaining food files are FDA-based. They are the Total Diet Study (TDS); the Food Labeling and Product Surveillance (FLAPS); and the Scientific Information Retrieval and Exchange Network (SIREN). More than 21,000 food products are indexed by Langual and searchable in the FMDB.

Nutritionists, food scientists and toxicologists have different information needs. Such questions as **What is the iron content of sugared breakfast flakes?** or **Do sugared breakfast flakes contain malathion?** may be answered by using Langual. Both questions may be answered by finding all **BREAKFAST CEREALS** that are **FLAKED** and **CONTAIN SUGAR**. Each of these elements is a retrieval term in Langual. Definitions explain that the "contains sugar" element is searched by using the index term **SUCROSE ADDED**.

When **BREAKFAST CEREAL** is entered into the Food Monitoring Database, food names from almost all the databases are retrieved. However, when the index terms **FLAKED** and **SUCROSE ADDED** are combined with the retrieval from **BREAKFAST CEREAL**, a much smaller number of food names is retrieved.

Example:

| # Food Names | Database |
|---|---|
| 29 | Nationwide Food Consumption Survey |
| 4 | SIREN |
| 2 | Total Diet Study |
| 20 | USDA Handbook 8 |
| 3 | FLAPS |
| 8 | Carotenoid Foods |

The names of the food products are then displayed. Each database is listed separately and each food name has a code identifier linking it to the database.

Example:

| | |
|---|---|
| USDA Handbook 8 | |
| Corn flakes cereal, low sodium | FOODID 08022 |
| Team cereal | FOODID 08075 |
| Total Diet Study | |
| Cereal, cornflakes | TDCODE 071 |
| Cereal, raisin bran | TDCODE 074 |

When food products sharing common characteristics are identified, the Food Monitoring Database search program is used to find nutritional or toxicological data. USDA Handbook 8, which contains results from nutritional analyses of foods, is used to answer the first question on the iron content of sugared breakfast flakes. The word **IRON** is entered into the FMDB and arrayed in an alphabetical list. It is then selected for retrieval and matched with the breakfast cereal food names found in Handbook 8. From the USDA nutritional data stored in the FMDB, a final report displays the results.

Example:
USDA HANDBOOK #8 REFERENCES

| ID # | Description | Mean # Obs. | Std. /100 g | Error |
|---|---|---|---|---|
| 08022 | Corn flakes cereal, low sodium | | | |
| | **IRON** | 2 | 2.22 | .177 |
| 08070 | Team cereal | | | |
| | **IRON** | 21 | 6.12 | .176 |

For the question on malathion the Total Diet Study would provide the answer. The word **MALATHION** is entered into FMDB and arrayed alphabetically. It is then selected and matched with the breakfast cereal names previously retrieved. A final report is displayed showing residue amounts of malathion in parts per million using TDS toxicological data.

Example:
TOTAL DIET STUDY REFERENCES

| TDCODE | Description | Basket Year | Residue Amount | Residue Unit |
|---|---|---|---|---|
| 071 | Cereal, cornflakes | 1981 | 0.002 | ppm |
| 074 | Cereal, raisin bran | 1985 | 0.700<br>0.006<br>0.015 | ppm |
| | | 1986 | 0.004<br>0.015 | ppm |

In summary, Langual not only provides a way of retrieving foods with common characteristics from a single database, but it also provides the capability for linking highly dissimilar databases through a universal retrieval language.

# Addressing and Related Issues in Electronic Mail Systems

**John C. Klensin***
**Lorry Scura+**

## Introduction

The worldwide "electronic mail network" is actually a collection of networks that do not have identical conventions for addresses or the way in which these are written. In several respects, the situation is not unlike that in the obtaining and use of food composition data. With the latter, one must first find the location of the data, then find out how to obtain it, then somehow translate its conventions, definitions, and coding and nomenclature methods into one's own. The good news about electronic mail addressing is that much more standardization-- both about the terminology used in description and the ways of accomplishing the translations-- than the situation we have with food composition data. Indeed, in many cases, a common-appearing address syntax is used, and transitions of electronic mail from one environment to another are completely transparent to the users. Nonetheless, some knowledge of alternate addressing models is useful, if only to facilitate intelligent asking for help when problems arise or unusual situations are encountered.

## Different Host Systems, Different Styles

In addition to actual addressing issues, "external" electronic mail has been developed on most systems as an extension of whatever "internal" electronic mail or interpersonal note facilities were available on those systems already. Some of these extensions have not been very neat because the addressing forms needed in "wide area" heterogeneous networks have been different from those previously developed for use in local systems. Again, in parallel with issues in the design of software for dietary analysis, there has been a tension between trying to make the interface look (to the user) similar to other, familiar, software in the local environment and trying to make network interface software look the same across all hosts on the network. The former usually wins out: the major advantages of the latter are convenience for the few of us who use the same networks from many different machines and the even fewer of us who try to write general addressing or electronic mail use guidelines or tutorials.

Two examples may illustrate the problems encountered by system designers, and passed on to users, when general network addressing is added to existing local electronic mail facilities.

- Some systems have historically been designed on the assumption that some few characters (typically about eight) are sufficient for representing the names of users and host machines. For a variety of reasons (a few discussed below), naming systems on the long-distance networks tend to require longer names, and more structure, than can be accommodated in two eight-character fields.
- Some systems "know", based on local conventions, what characters can appear in host names and people's names or other addresses. Often these assumptions exclude characters such as period ("."), at-sign ("@"), semicolon (";"), or slash ("/"), which are heavily used in various national and international network addressing systems.

The combination of these two problems results in some odd-appearing local syntax and other arrangements: the need to send mail to intermediate (e.g., eight-character or restricted syntax) addresses with internal instructions about the "real" address or to use quoted strings or special processing instructions to permit processing of locally-unacceptable addresses. Fortunately, over time, enough host operating system work seems to be going on that the number and complexity of these "work arounds" that are required is decreasing. For example, many VAX/VMS systems use a syntax that is illustrated in some of the slides for our other paper. There, Internet or BITNET addresses must be prefaced by, e.g., 'IN%"' and followed by another double-quote, to avoid conflicts with local parsing conventions. This is typical of what is done to work around local systems and their conventions.

That said, these arrangements, whether odd typing conventions or intermediate addressing is required, are just a nuisance: they need to be understood and worked out with local support personnel and then observed as a ritual without spending much further energy worrying about just what they mean.

**The Networks and Their Connections**

Within the US, there are five major network arrangements that can be used for electronic mail. These are interconnected, although some of the connections are easier to use than others. Hosts and other arrangements that provide connections and translations between networks are known as "gateways". Most of the rest of the world that has network connectivity is linked to one or more of these, sometimes under different names and often with less complicated arrangements. In no particular order, they are:

**BITNET:** An academic and research network that primarily connects to academic institutions although some government components and private research organizations are also members. "Native" BITNET addresses are of the form username@hostname, with the user and host names each being limited to eight characters. Many BITNET hosts are converting, or have converted, to Internet (see below) naming conventions, hence "UCBCMSA" (sometimes, by convention, called "UCBCMSA.BITNET") and "CSMA.BERKELEY.EDU" are the same machine and can be addressed either way. From the Internet, BITNET users are addressed either in username@domain form, where "domain" is the Internet name for the host, or using the form user%host.bitnet@gateway. When no better gateway is known, cunyvm.cuny.edu is typically used.

BITNET users at most hosts send mail to Internet addresses simply by specifying the Internet address; the exceptions are becoming less and less useful as hosts. As Internet connections become more widespread, it is possible that BITNET service will be phased out entirely during the next few years. BITNET itself is run by an organization called CREN, which is US-based. A European network called EARN, one in Canada called NETNorth, and a few smaller ones in other parts of the words are administratively separate from, but technologically part of, BITNET.

**Internet:** This is actually a collection of networks running the same protocols and sufficiently interconnected so, to the user, they appear to be a single network. Its components include the so-called "NSFNet regional networks" (that terminology may be obsolete by the time you read this) such as BARRNet, CERFNet, NYSERNet, JvNCNet, NEARNet, and so on. It evolved from the old ARPANET and is expected to evolve gradually into a very high performance "National Research and Education Network" (NREN). It is practical to connect much smaller machines (e.g., individual workstations or personal computers) to the Internet than has historically been the case with BITNET, so the number of hosts has become very large, with thousands of them at some universities. Although rules and arrangements differ from one region to the next, the Internet serves a superset of the institutions that would be accepted for BITNET membership. Internet addressing takes the form user@domain, where either the user name or the domain name may be quite long. Some Internet sites actually use full personal names, such that John_C_Klensin@MIT.EDU would be a plausible address. Domain names use a hierarchical structured model to identify a particular host. For example, in the address Lorry_SY@nutmeg.hnrc.tufts.edu, "nutmeg" is a host at the HNRC, which is part of Tufts, which is an EDUcational institution. Internet addresses outside the US typically use the two-letter country abbreviation as the last component, e.g., tubvm.cs.tu-berlin.de identifies a machine associated with the Technical University of Berlin in Germany. For academic and research network purposes, the Internet may be considered the "spine" through which most other connections flow.

**UUCP and FIDONET:** These two names refer to protocol arrangements that have been used to create networks connecting UNIX (tm) machines and things that imitate them and IBM PC-like desktop computers and things that imitate them, respectively. They, especially the latter, have been excellent solutions for low-cost, minimum fuss connections for individuals and small companies. Both networks have become extensive, and provide the major or only connections to some countries. Both user Internet-style addressing, and gateways to other networks that use this style of addressing are typically invisible to the end-user.

**Commercial providers:** There are several commercial providers of electronic mail services. Despite using "network" terminology, most of them operate with a single tightly-connected cluster of hosts and, except for connections to other networks, essentially accept and deliver mail on the same machine. The best-known among these are probably ATTMAIL, CompuServe, EasyLink, MCIMail, and various services offered by BT (e.g., DIALCOM, CGNET) and US Sprint (e.g., SprintMail) (these are all trademarks of their respective owners). All of these have connections to the Internet, usually using standard Internet addressing with some prefix that designates the network. In general, these commercial providers will sell electronic mail services (and other computer, teleconferencing, and bulletin board services) to anyone who is willing to pay the bills; several of the services are remarkably inexpensive for relatively light use. The much-advertised GENIE and Prodigy do not offer external mail connections at this time. Many of these systems are beginning to use so-called X.400 addressing: insofar as the Internet's domain names represent a single hierarchy, X.400 is multi-hierarchical, with separate hierarchies for various organizational arrangements and types of identification. There is, as yet, no Standard way to write an X.400 address; the syntax differs from one system to another.

## Getting Attachment Help

The best advice is to find an expert and ask. If the advice you get seems to contradict good sense or what appears above, get a second opinion. Most academic institutions in the US are connected to either BITNET, the Internet, or both, although these networks are often thought of as the property of computer and physical scientists until someone else asks. Individual practitioners are likely to find one of the commercial services attractive, especially if there are few computer resources in-house. Most of those services offer a mix of databases, fax and telex arrangements, and user support services in addition to their electronic mail and comparison shopping is sensible. UUCP and Fidonet connections can be very attractive and inexpensive once established, but can be difficult to set up initially unless one has done so before.

## Electronic Mail and the Nutrient Composition Community

In discussing why one would want to use these facilities, we pointed out that topic-specific electronic mailing lists may be much more useful for some purposes than person-to-person electronic mail (or postal mail, telephone, or fax). Electronic mailing lists are especially useful for asking a question of, or discussing issues among, a group of people. INFOODS has created such a list for the nutrient composition community, paralleling the European WHO list for Nutritional Epidemiology. The instructions for subscription to the two lists (one may, of course, subscribe to either or both) once you have electronic mail access operating are:

### *Nutritional Epidemiology List*

Send a mail message whose text body contains the line

SUB NUTEPI your full name

The mail message should be addressed to:

From BITNET or EARN: LISTSERV@db0tuil1

From other networks, use the Internet address form:

LISTSERV@tubvm.cs.tu-berlin.de

Once you are on the list, administrative requests (e.g., subscription and subscription-cancellations) should be sent to the LISTSERV, as shown above. Material to be posted to other list subscribers should be sent to

NUTEPI@db0tuil1 or
NUTEPI@tubvm.cs.tu-berlin.de respectively.

### *Food Composition Discussion List*

Send a mail message requesting a subscription to:

food-comp-request@infoods.mit.edu

This mail is read by a person, so the message need not be in a specific form. The address is the same from BITNET, EARN, the Internet, etc.

Once you are on the list, administrative requests (e.g., subscription and subscription-cancellations) should be sent to the "-request" address, as shown above. Material to be posted to other list subscribers should be sent to

food-comp@infoods.mit.edu

## References and Useful Addresses

The following information was accurate as of the time of the conference. Some of this information changes very rapidly and, in particular, it has been claimed that any description of networks and connectivity is always obsolete by the time it can be published.

Frey, Donnalyn and Rick Adams. *!%@:: A Directory of Electronic Mail Addresses*. O'Reilly and Associates, Petaluma, CA: 1989.

Quarterman, John S. *The Matrix, Computer Networks and Conferencing Systems Worldwide.* Digital Press, 1990.

The vendors of commercial electronic mail systems are usually listed in local "white pages" telephone books. The following information may be useful for those that often are not:

CompuServe starter kits are available from most computer dealers. Telephone enrollment and more information is available at 800-

MCIMail registers new users and provides additional information by telephone at 800-444-6245 (202-833-8484 in the Washington, DC area).

SprintNet (SprintMail, Telemail) customer service and registration information is available at 800-336-0437.

*John C. Klensin
INFOODS Secretariat, Room N52-457, Massachusetts Institute of Technology, 77 Massachusetts Avenue, Cambridge, MA 02139.

+Lorry Scura
USDA Human Nutrition Research Center, Tufts University, 711 Washington St., Boston, MA 02111.

# Food Frequency Data Entry and Analysis Program

**R. Sue McPherson**
**University of Texas**
**Houston, Texas**

The Food Frequency Data Entry and Analysis Program (FFDEAP) is designed to assist in the development, data entry and nutrient analysis of frequency of food use data. The FFDEAP is not accompanied by a food and nutrient data base. Because of the desirability that the foods included on a food frequency questionnaire be study specific, and that, in some cases, a number of similar foods may be aggregated into a generic food type (e.g., citrus fruits, breads, crackers, etc.), the user can select both the foods that will appear on the food frequency questionnaire and the corresponding gram weights (portion size information) and nutrient information that will be used in the nutrient analysis. The FFDEAP gives the user the unique capability of developing, collecting and analyzing frequency of food use information specific to each study population or the goals of a project.

## FOODS AND NUTRIENTS

In order to operate the program, the user must develop two files. The user manual contains detailed instructions and file formats for these two files. The first user-developed file, the GRAMS file, will contain the food and gram weight information. The foods in the GRAMS file are those that the user has selected for inclusion on the food frequency questionnaire. The food list cannot exceed 250 items. This file also contains the gram weight of the selected food model and portion size options for each food. The second user-developed file, the NUTRIENT file, contains the nutrient information per 100 grams of each food on the food list. The number of nutrients for each food cannot exceed 50.

The gram weight and nutrient information for the foods can be obtained from any source, including nutrient data bases, the literature or from laboratory analyses. Users have the freedom to select their source of nutrient information.

## ENTRY AND EDITING

The FFDEAP has two parts, the ENTRY Program for data entry and the ANALYSIS Program for nutrient analysis and reporting. In the ENTRY program, subject information data can be entered into the Cover Sheet Information Screen. The Subject Number refers to the unique subject identification number. The Date distinguishes the date of the food frequency interview. Other identifying data entry fields are available on the Cover Sheet Information Screen to aid in study management and quality control. The ENTRY Program reads the user-created GRAMS file and presents the foods on the screen one by one in the order in which they appear on the food frequency questionnaire. The ENTRY Program has on-line editing capabilities to allow additions, deletions, or changes to the data entered. A sample of a FOOD FREQUENCY DATA ENTRY SCREEN follows:

```
FIELD PGUP/PGDN-ITEM ^HOME-TOP FILE ^END-END FILE
ESC to Exit

  SUBJECT #: 1          DATE: 12/09/90

 Food: 1 WHOLE MILK

 Eats Food: (1) No (2) Yes (9) Unknown: 2

 Enter Frequency: Monthly [ ]   Weekly [ ]   Daily [ ]

 Weight (GM,OZ,LB,MG,KG)        Fluid (FO,PT,QT,GAL,ML,L)
 Cup (CP,C)                 Spoon (TS,HTS,TB,HTB)
 Glass (G1-G5)              Bowl (B1,B2)
 Mug (MG1,MG2,MG3)          Serving (SV)
 Not Further Specified (NFS)   Unknown (UNK)

        Usual Portion Size From Food Models
 How Many:  1     Food Model Name: [          ]
```

The FFDEAP is primarily designed to collect and analyze data for a 28-day time period. However, data can be collected for shorter periods of time and entered into the FFDEAP using only the weekly and daily data entry fields.

Nutrient analysis of the entered data is carried out in the ANALYSIS Program of the FFDEAP. In the ANALYSIS Program the user may develop and edit printed report forms and determine which nutrients will be in the reports and ASCII files.

### PORTION SIZE OPTIONS

For each food on the food frequency questionnaire, the user can determine the food models and portion sizes that will be allowed for that food. The portion size information is then entered into the GRAMS file and becomes part of the quality control data for the data entry. The allowed food models and portion sizes for each food will also be displayed on the food frequency data entry screen in the center box as shown in the above screen. If a food frequency questionnaire is designed to be semi-quantitative, the information for one standard portion for each food can be entered.

The food model and portion size options can be household measures such as cups and tablespoons, weight and fluid measures such as grams, ounces and fluid ounces or two-dimensional geometric models. The two-dimensional geometric models are provided with the manual and include rectangles, circles, wedges and two-dimensional representations of mugs, cups, glasses, spoons, bowls, and mounds. Even though all these food model and portion size options are allowed and recognized by the FFDEAP, the user may elect to use a limited number of these available.

### CUSTOMIZED REPORTS AND ASCII FILES

The ANALYSIS Program allows the user to create printed reports and/or ASCII files. The user may select the nutrients to be included in the ASCII files independently from those to be included in the report. Reports can be sent to a file and printed at a later time.

The Report is available in both detailed and summary format:

(1) The summary portion of the report contains the total nutrient intake for the month as well as an average daily nutrient intake based on a 28-day month.

(2) The detailed portion of the report also contains a detailed nutrient profile for each food. The grams per month consumed for each food is reported, and the nutrients based on the monthly intake of each food. Reports can be tailored to contain only the nutrients of interest to the user. A sample report is attached.

Two ASCII files are produced by the ANALYSIS Program which can be read easily by data base management, statistical, and other software programs for aggregation and summarization of data.

### THE FFDEAP PACKAGE

The PROGRAM DISK contains the FFDEAP Files.

The SAMPLE DISK contains examples of the files generated by the FFDEAP as well as examples of the GRAMS and NUTRIENT files.

The USER MANUAL contains detailed instructions on developing the GRAMS and NUTRIENT files, on operating the ENTRY and ANALYSIS Programs of the FFDEAP, and a comprehensive TUTORIAL.

The SAMPLE FOOD FREQUENCY FORM includes a cover sheet and is designed to correspond to the data entry fields on the screen. Users can use it for data collection or as a guide for developing their own forms.

The TWO-DIMENSIONAL FOOD MODELS are representations of three dimensional geometric shapes and common household measures. Examples of the types of shapes represented include circles, rectangles, and wedges. Examples of the common household measures include bowls, glasses, mugs, spoons and cups. The TWO-DIMENSIONAL FOOD MODELS included with the package are camera ready and can be duplicated, laminated, bound, etc. for use in studies.

# Nutrient Databases for Food Frequency Instruments

**Laura Sampson**
**Harvard University**
**Boston, Massachusetts**

## ABSTRACT

The food frequency questionnaire (FFQ), as a result of continued development beginning in the 1950's, has become the main-stay for nutritional epidemiology for a variety of reasons. The FFQ consists of two basic parts: a food list and frequency responses. The creation of a FFQ should involve pretesting to design a form specific to the target population. FFQ's can be reasonably reproducible and valid for both specific foods and nutrients. Testing of the FFQ, especially at the food level, provides useful insight into the ability of the form to provide adequate dietary information. The tool may be adapted to obtain dietary information more efficiently under a variety of situations. The FFQ has been used to successfully predict disease.

This presentation will briefly review the development of Food Frequency Questionnaire's (FFQ's) and more extensively describe the development and refinement of the Harvard/Willett food frequency questionnaire. With the results of validation studies conducted over the last 10 years, the form has been continuously refined to improve its quality and has been used to test numerous hypotheses relating diet and disease. Two recent studies clearly demonstrate that FFQs are capable of prospectively relating diet to disease. This work substantiates why FFQ's have become a mainstay for nutritional epidemiology.

The origins of the FFQ occurred in 1940's when Burke developed a detailed dietary history interview that evaluated an individual's usual diet through three methods: a 24 hour recall, a 3 day diet record, and a checklist of foods consumed during the past month(1). This checklist was apparently the first example of a FFQ, and was further developed in the 50's by Stefanik and Trulson, and others as outlined on the slide (2-5). At this time the American Public Health Association through a committe report called for investigators to

| Investigator | Year | Tool |
|---|---|---|
| Burke | 1947 | Detailed History |
| Stephanik & Trulson | 1962 | FFQ |
| Heady | 1961 | FFQ |
| Weihl & Reed | 1960 | FFQ |
| Marr | 1971 | FFQ |

find simple methods by which dietary data could be obtained quickly and inexpensively to provide nutritional information on a large number of subjects. This was deemed necessary to improve insight into the ever growing public health concerns in our country -- cancer and heart disease. In the early seventies there was skepticism concerning the use of FFQ's in predicting disease but with improved questionnaire formats and more studies reporting the usefulness of the form, interest has returned (6). FFQ's have now become widely accepted as a tool by which nutritional information can be obtained from large numbers of subjects necessary to investigate diet and disease risks.

Our work began in the late 70's under the direction of Walter Willett and with the help of Jelia Witschi, a registered dietitian from the Harvard School of Public Health. A FFQ consists of 2 basic parts: (1) a list of foods and (2) a set of frequency response options used to indicate how often each food is consumed during a specified period of time. The objective is to rank individuals according to food and nutrient intake and to characterize food intake over an extended period. Usual intake is more relevant when studying diet and most diseases than today's or yesterday's food intake. To accomplish this, a long list of foods that contained significant amounts of the nutrients of interest was created. Second, a small scale pilot study was completed to assist in identifying infrequently eaten foods, which were eliminated, bringing the number of foods to 99. Third, the form was pretested to see which foods best predicted the nutrients of interest. This 99-item form was mailed to 2000 randomly selected Nurses Health Study subjects. 87% returned the form. Nutrient intake calculated from the form was used as the dependent variable in a step-wise multiple regression equation to determine which foods best predicted the nutrients of interest.

**Vitamin C Intake**

| Foods Most Predictive of Between Person Variation | Cumulative $R^2$ |
|---|---|
| Supplements | 0.91 |
| Orange Juice | 0.93 |
| Multivitamins | 0.94 |
| Fruit Punch | 0.96 |
| Spinach, Greens | 0.96 |
| Berries | 0.96 |
| Brussel Sprouts | 0.96 |

For example, 96% of the between-person variance, what one person eats versus another, in Vitamin C intake is explained by 7 foods. As a result of the compilation of this data, it was determined that 61 foods plus multivitamin usage accounted for at least 80% of the between-person variation in the specified 18 nutrients analyzed.

The 1980 FFQ that was distributed to over 100,000 women between the ages of 34 to 50 years in 11 US states contained 61 foods separated into 6 groups - dairy; fruits; vegetables; meats; sweets, baked goods, and cereals (as 1 group); and a miscellaneous group. Each food was assigned a specified portion using natural units when possible (ie slice of bread, 1 egg, glass of milk, 1 apple, 1 pat of margarine). The subjects were also supplied with 9 frequency responses by which to identify how often they ate the specified portion of food. The frequency responses range from never to 6 plus per day. This allows for subjects to answer how often they eat infrequent foods like liver as well as margarine or butter. The subjects were asked to describe their usual intake of these foods over the past year. There were also several special features of the form. The subjects were asked to identify the type of fat used for frying and cooking, the type of margarine and cereal, and amounts of bran, and sugar added to foods. Also, multivitamin and vitamin A,C and E use was obtained.

This FFQ provided our first large dietary data set, and was the object of our initial reproducibility and validity studies. Over the subsequent 10 years, other forms representing expansion, refinements, and modifications for specific populations have been created. There are 12 versions today and it has been translated into Spanish for Mexican and Latin American studies, French, Italian, and Greek.

What did we find?

Data obtained by a FFQ can be used in several ways. The actual frequency or number of times a food or food group is consumed might be of interest. Another is the nutrient scores calculated by summing the total of the product of the frequency of consumption of the foods by their nutritional content of a specified or assumed portion.

The first tests of the usefulness of the 1980 FFQ were done at the nutrient level as part of a substudy of the large prospective study mentioned. Reproducibility studies designed to see how well the form per-

3. For each food listed, fill in the circle indicating how often on average you have used the amount specified during the past year.

| DAIRY FOODS | AVERAGE USE LAST YEAR: Never, or less than once per month | 1-3 per mo. | 1 per week | 2-4 per week | 5-6 per week | 1 per day | 2-3 per day | 4-5 per day | 6+ per day |
|---|---|---|---|---|---|---|---|---|---|
| Skim or low fat milk (8 oz glass) | ○ | ○ | Ⓦ | ○ | ○ | Ⓓ | ○ | ○ | ○ |
| Whole milk (8 oz glass) | ○ | ○ | Ⓦ | ○ | ○ | Ⓓ | ○ | ○ | ○ |
| Cream, e.g. coffee, whipped (Tbs) | ○ | ○ | Ⓦ | ○ | ○ | Ⓓ | ○ | ○ | ○ |
| Sour cream (Tbs) | ○ | ○ | Ⓦ | ○ | ○ | Ⓓ | ○ | ○ | ○ |
| Non-dairy coffee whitener (tsp) | ○ | ○ | Ⓦ | ○ | ○ | Ⓓ | ○ | ○ | ○ |
| Sherbet or ice milk (½ cup) | ○ | ○ | Ⓦ | ○ | ○ | Ⓓ | ○ | ○ | ○ |
| Ice cream (½ cup) | ○ | ○ | Ⓦ | ○ | ○ | Ⓓ | ○ | ○ | ○ |
| Yogurt (1 cup) | ○ | ○ | Ⓦ | ○ | ○ | Ⓓ | ○ | ○ | ○ |
| Cottage or ricotta cheese (½ cup) | ○ | ○ | Ⓦ | ○ | ○ | Ⓓ | ○ | ○ | ○ |
| Cream cheese (1 oz) | ○ | ○ | Ⓦ | ○ | ○ | Ⓓ | ○ | ○ | ○ |
| Other cheese, e.g. American, cheddar, etc., plain or as part of a dish (1 slice or 1 oz serving) | ○ | ○ | Ⓦ | ○ | ○ | Ⓓ | ○ | ○ | ○ |
| Margarine (pat) added to food or bread, exclude use in cooking | ○ | ○ | Ⓦ | ○ | ○ | Ⓓ | ○ | ○ | ○ |
| Butter (pat), added to food or bread, exclude use in cooking | ○ | ○ | Ⓦ | ○ | ○ | Ⓓ | ○ | ○ | ○ |

**Summed total daily cholesterol calculations**

| | Nurses Scores | Weighted Daily Freq. | x | Nutrient Portion |
|---|---|---|---|---|
| Eggs (1/day) | 274.0 mg = | 1 | x | 274.0 |
| Milk (2-3/day) | 82.8 mg = | 2.5 | x | 33.0 |
| Ice Cream (5-6/weeks) | 23.6 mg = | .8 | x | 29.5 |
| | 380.1 mg = | Summed Total Daily Cholesterol | | |

formed when it was repeated, and the validation studies, designed to see how well the form did when it was compared to other dietary methods, were conducted (7). 28 days of diet records were collected from 194 subjects in the Greater Boston Area. The days were spaced evenly throughout the year to acquire 7 days from each of the seasons. At the end of the year of diet record keeping the FFQ was readministered to the women.

The results of these studies are shown on the next few slides. Pearson correlation coefficients for a reliability study, comparing Questionnaire 1 to Questionnaire 2 (Q2) ranged from 0.52 for Vitamin A without supplements to 0.71 for sucrose and produced a mean of 0.61 for 15 nutrients. Correlation coefficients of nutrient reproducibility studies are often between 0.5 and 0.7. Though these values may seem low compared to measurements made in a controlled laboratory enviroment they are comparable to biological measurements such as serum cholesterol or blood pressure taken over time.

The validity correlations shown here report the comparison of Q2 which reflects the average nutrient intake over the last year, the year of diet record keeping, with the mean nutrient intake from the diet records. They ranged from 0.36 for Vitamin A to 0.75 for Vitamin C . When we ranked individuals intake from the FFQ and the diet records into quintiles only 3% were extremely misclassified. This clearly shows the ability of the form to rank individuals into high versus low intake groups, the primary objective of nutritional epidemiologists. Once subjects are separated into high versus low intake groups their risk to a particular disease can be assessed.

What did we learn and how did we modify the form?

First, the form was converted to an optically scannable form. Our original form was a self administered mailed form which required key punching. Optical scanning decreased our coding and data entry costs.

Secondly, the form was expanded from 61 to 116 food items (8). Foods that had been collapsed or grouped together on the form were itemized separately to provide the subjects with more simple clear questions. We added back some foods that were important sources of nutrients of interest even though they did not independently contribute to between-person variation in nutrient intake in stepwise multiple regression. Four foods were added after analysis of the food records showed that they contained sources of nutrients that contributed to the absolute intake of the population. They were mixed vegetables, tomato sauce, other fruit juices and English muffins, bagels, and rolls as one group.

The expanded form was completed in 1984 four years after the first by the same subjects that had completed the original FFQ and diet records. The 61 item form had accounted for 51-97% of the nutrients consumed in the diet records. The expanded version accounted for 86-99%. Correlations comparing the revised and the original questionnaire completed by the Boston area subjects four years later ranged from 0.44 for total carbohydrate to 0.62 for vitamin C including supplements. This shows that the FFQ can provide useful information about nutrient intakes in the past.

**Percentage of Intake Accounted for by Foods on FFQ's**

| *Nutrient* | *61-Item FFQ* | *116-Item FFQ* |
|---|---|---|
| Total Calories | 69 | 93 |
| Protein | 77 | 95 |
| Total fat | 70 | 96 |
| Saturated fat | 75 | 96 |
| Monounsaturated fat | 72 | 96 |
| Cholesterol | 85 | 97 |
| Total carbohydrate | 61 | 90 |
| Crude fiber | 64 | 86 |
| Sucrose | 78 | 92 |
| Total vitamin A | 77 | 96 |
| Vitamin C | 84 | 93 |
| Vitamin B1 | 81 | 95 |
| Vitamin B2 | 85 | 95 |
| Vitamin B6 | 97 | 99 |
| Calcium | 77 | 94 |
| Iron | 75 | 93 |
| Mean | 75 | 94 |

Thirdly, an open-ended section to the questionnaire where subjects could report foods that were eaten weekly that were not asked on the FFQ was added. It was important to evaluate the contribution not only of the new additional foods section but also the other open-ended questions stated earlier, such as the types of oil, cereal, and multivitamin usage primarily because open-ended sections require additional coding steps and increase costs. This was done at the same time as the study just mentioned where the FFQ's were completed four years later. To investigate this question, nutrient intakes were computed from the FFQ's with and without the open-ended questions being included into the analysis and were compared to the diet record intakes. Only modest increases in correlations due to the addition of the open-ended questions in some of the nutrients were found. None of the changes were statistically significant. Although these results show little effect in the estimation of nutrient intake and the open-ended questions are costly and time consuming to process, caution should be taken when considering the removal of these questions from the form. These sections may provide important information in a more complex or heterogeneous population. A major limitation of FFQ's is that the food listing is fixed so the option for subjects to provide additional information maybe well worth the cost.

**Pearson Correlations between diet records and 116-Item FFQ**

| *Nutrient* | *No Deletion* | *Open-ended sections deleted* |
|---|---|---|
| Protein | 0.52 | 0.53 |
| Total fat | 0.54 | 0.54 |
| Saturated fat | 0.52 | 0.52 |
| Polyunsaturated fat | 0.58 | 0.58 |
| Monounsaturated fat | 0.48 | 0.47 |
| Cholesterol | 0.57 | 0.57 |
| Total carbohydrate | 0.61 | 0.61 |
| Crude fiber | 0.56 | 0.55 |
| Sucrose | 0.45 | 0.45 |
| Total Vitamin A | 0.44 | 0.38 |
| Vitamin C | 0.54 | 0.53 |
| Vitamin B1 | 0.58 | 0.52 |
| Calcium | 0.56 | 0.50 |
| Phosphorus | 0.51 | 0.50 |
| Potassium | 0.53 | 0.53 |
| Iron | 0.55 | 0.42 |

Fourthly, the controversial issue of portion sizes on the form was addressed.The Harvard/Willett FFQ is termed a semiquantitative FFQ because it specifies typical portion sizes. A pure FFQ would not request subjects to provide information on portion sizes. Other forms of FFQ's ask subjects to estimate their usual portion sizes. To look at this issue, the number of times that 66 food were eaten and the portion size on each occasion during the four weeks of diet records in the Boston area subjects were determined (9). For each food, total population variance in portion size was separated into within- and between-person components. A large within-person variance would indicate that subjects do not consistently eat the same portions and a large between-person variance would indicate that an individual might eat the same portion but that portion is not eaten often by other people. For all but 7 food items the within-person (intraindividual) variation was higher than the between-person (interindividual) variation. Foods with a high within-person variation also tended to have a high between-person variance. This data suggests that describing a "usual" portion size maybe a difficult thing for subjects to do since they do not eat the same portion. A portion size that at least approximates a population's norm identified through pilot testing at least will add clarity to the food frequency question. It will require that subjects calculate frequencies based on the specified portion and this may add complexity too. But adding the portion information allows for quantitative food and nutrient data to be obtained.

In a general population sample involving 97 subjects nutrient intakes derived from a FFQ without serving sizes for foods that do not come in natural units and from the same FFQ plus portion size information obtained during an extensive interview were both compared with intakes calculated from 7 day diet records (10). Correlations of log nutrient intakes for calories and macronutrients showed very slight increases when portion size information was used. After adjustment for total energy intake, there was essentially no improvement using the additional portion size data. There was no pattern of improvement for micronutrients. These data suggest that additional portion size information does not necessarily increase the validity of estimates obtained from a SFFQ alone.

The results reported above were based on reproducibility and validity studies at the nutrient level.We have also completed reliability and validity studies of the FFQ at the food level. To evaluate reproducibility at the food level, we compared the 99 food-item pilot FFQ completed by 1497 subjects with their responses on the 61 item FFQ that was completed 9 months later (11). Correlations were highest for beverages with a correlation of 0.7 and ranged from 0.6 - 0.7 for foods eaten frequently and from 0.34-0.45 for foods eaten infrequently. In the 1980 study when the first and second FFQ were compared the reproducibility correlations ranged from 0.24 for fruit punch to 0.93 for beer. The mean correlation was 0.57. For 23% of the food items the correlation coefficient was greater or equal to 0.70, and for 73% of the foods it was greater or equal to 0.5. People tend to report well what they eat often and less well what they eat less often.

To evaluate validity at the food level, the original FFQ and the diet records were compared. To do this, the foods coded from the diet records were grouped to correspond with a food on the questionnaire (12). For example 146 diet record meat codes represented the two meat items on the FFQ, 10 codes were compressed to correspond with skim or lowfat milk . The amounts of the foods reported in the diet records were converted into the amounts specified on the questionnaire. For each subject we determined the mean daily amount of food consumed during each of the 4 weeks of diet records and the mean daily amount for the 4 full weeks of recording. Pearson correlation coefficients between the second FFQ and the diet records corrected for within- versus between-person variance ranged from 0.17 for yellow squash to 0.94 for beer. Subjects tended to overestimate consumption of some foods on the FFQ. In comparing the absolute amounts of each food estimated by the FFQ and the diet records it was found that butter, whole milk, eggs, processed meats and cold breakfast cereal were underestimated by 10-30% on the FFQ. On the other hand certain foods were overestimated by at least 50% They consisted of a number of fruits and vegetables, skimmed or lowfat milk, yogurt, and fish. This suggests that people over-report consumption of "healthy" foods and underestimate those considered "undesirable".

In general this level of reproducibility and validity testing at the food item level shows the ability of the

| Underestimated 30% | Overestimated by 10 - by at least 50% |
|---|---|
| Butter | Some Fruits and Vegtables |
| Whole Milk | Skimmed or Lowfat Milk |
| Eggs | Yogurt |
| Processed Meat | Fish |
| Cold Breakfast Cereals | |

FFQ to document and quantify intakes. It also allows insight to be gained into why certain foods or nutrients have lower or higher correlations than others. This methodology will not be reviewed here. However, for example, from investigating the form at the food level we have separated the spinach question on the form into 2 separate items, and have separated carrots into separate raw and cooked questions.

In 1986 we reevaluated the reproducibility and validity of our expanded and refined FFQ which incorporated all of our changes to improve clarity of the form. The FFQ was evaluated in a group of 127 men and 197 women from the Greater Boston area. The women's data is not available at this time. For the men's study reproducibility of the FFQ completed one year apart ranged from 0.56 for polyunsaturated fat to 0.80 for vitamin C (13). The de-attenuated correlation coefficient or rather those corrected for the error due to within-person variability in diet records, between the questionnaire and the diet records ranged from 0.37 for polyunsaturated fat to 0.92 for vitamin C. The average correlation was 0.68.

As has been stated, a primary purpose of the FFQ is to be able to rank individuals into high and low intake groups to predict disease risks. It is best to evaluate this capacity in prospective studies to avoid the bias that might exist in case control studies. Even with prospective studies, caution must be used in interpreting these results because failure to find a relationship may simply be due to an inappropriate interval of time between FFQ's and the diagnosis of the disease. Absence of an expected association should not automatically be interpreted as the fault of the form.

We have conducted 2 studies which validate the use of the FFQ at the disease level by predicting strongly suspected associations. In 1980 when diet information was gathered from the 100,000 women, medical information was also obtained which has allowed prospective documentation of diet and disease risks. By 1986, 150 cases of colon cancer were documented (14). After adjusting for total energy intake, animal fat was positively associated with the risk of colon cancer. When subjects were separated into quintiles the relative risk for the highest as compared to the lowest group was 1.89.

In a second prospective study 52,000 men completed a 131 item FFQ in 1986. 7,248 participants underwent colonoscopy or sigmoidoscopy during a 2 year period. After adjusting for total caloric intake, red meat and saturated fat were positively associated with risk of adenomas of the large bowel and fiber intake from vegetables, fruits, and grains was strongly related to decreased risk (15).

In conclusion, the FFQ is an established nutritional tool that can be used with large numbers of subjects to acquire reproducible and valid food and nutrient information to test disease risks.

## References


Burke, B.S. The dietary history as a tool in research. J.Am. Diet. Assoc. 1947; 23:1041-1-46.

Stefanik, P.A., Trulson. M.F. Determining the frequency of foods in large group studies. Am. J. Clin. Nutr. 1962; 11:335-343.

Heady, J.H. Diets of bank clerks: Development of a method of classifying the diets od individuals for use in epidemiologic studies. J.R. Stat. Soc. (A) 1961; 124:336-361.

Wiehl, D.G., Reed, R. Development of new or improved dietary methods for epidemiological investigations. Am. J. Pub. Health 1960; 50:824-828.

Marr, J.W. Individual dietary surveys: Purposes and methods. World Rev. Nutr. Diet 1971; 13:105-164.

Willett WC. Nutritional Epidemiology. Oxford University Press, N.Y. 1989 (see Chapter 6).

Willett WC, Sampson, L, Stampfer MJ, Rosner B, Bain C, Witschi J, Hennekens CH, Speizer, FE. Reproducibility and validity of a semiquantitative food frequency questionnaire. Am J Epidemiol 1985; 122:51-65.

Willett W, Sampson L, Browne ML, Stampfer MJ, Rosner B, Hennekens CH, Speizer FE. The use of a self-administered questionnaire to assess diet four years in the past. Am J Epidemiol 1988; 127:188-199.

Hunter DJ, Sampson L, Stampfer MJ, Colditz GA, Rosner B, Willett WC. Variability of portion sizes of commonly consumed foods among a population of U.S. women. Am J Epidemiol 1988; 127:1140-1149.

Hernandez-Avila, M., Masters, C., Hunter, D., Buring, J., Phillips, J., Willett, W.C. Influence of additional portion size data on the validity of a semi-quantitative food frequency questionnaire. Am. J. Epidemiol. 128, 891.

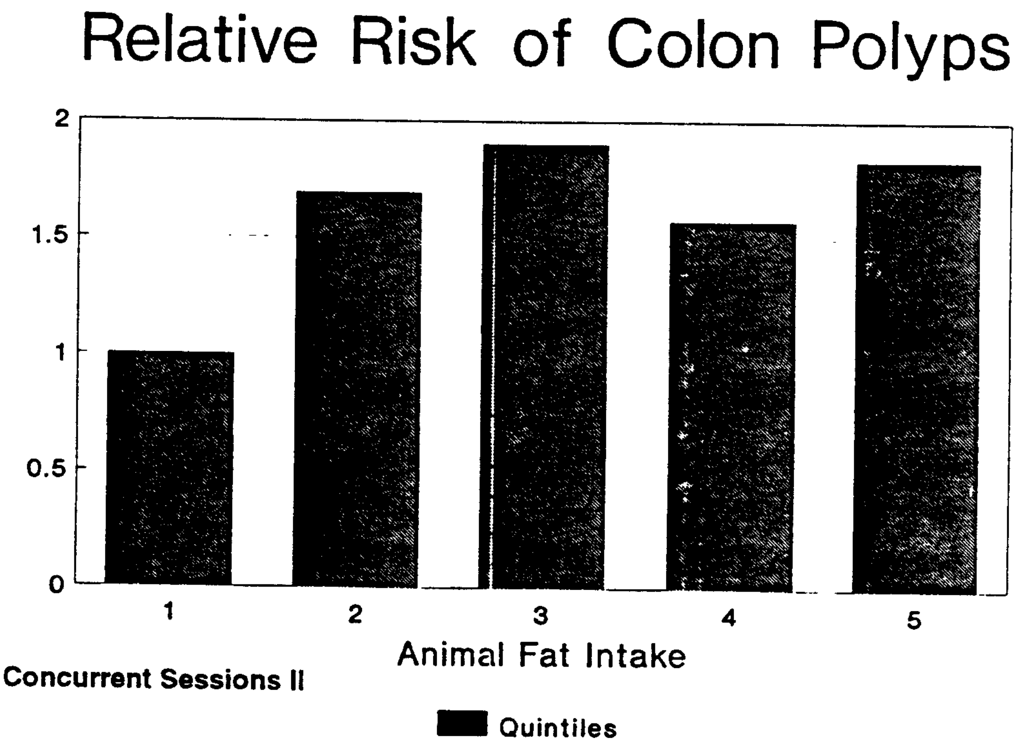
Colditz GA, Willett WC, Stampfer MJ, Sampson L, Rosner B, Hennekens CH, Speizer FE. The influence of age, relative weight, smoking, and alcohol intake on the reproducibility of a dietary questionnaire. Int J Epidemiol 1987; 16:392-398.

Salvini S, Hunter DJ, Sampson L, Stampfer MJ, Colditz GA, Rosner B, Willett WC. Food-based validation of a dietary questionnaire: the effects of week-to-week variation in food consumption. Int J Epidemiol 1989; 18:858-67.

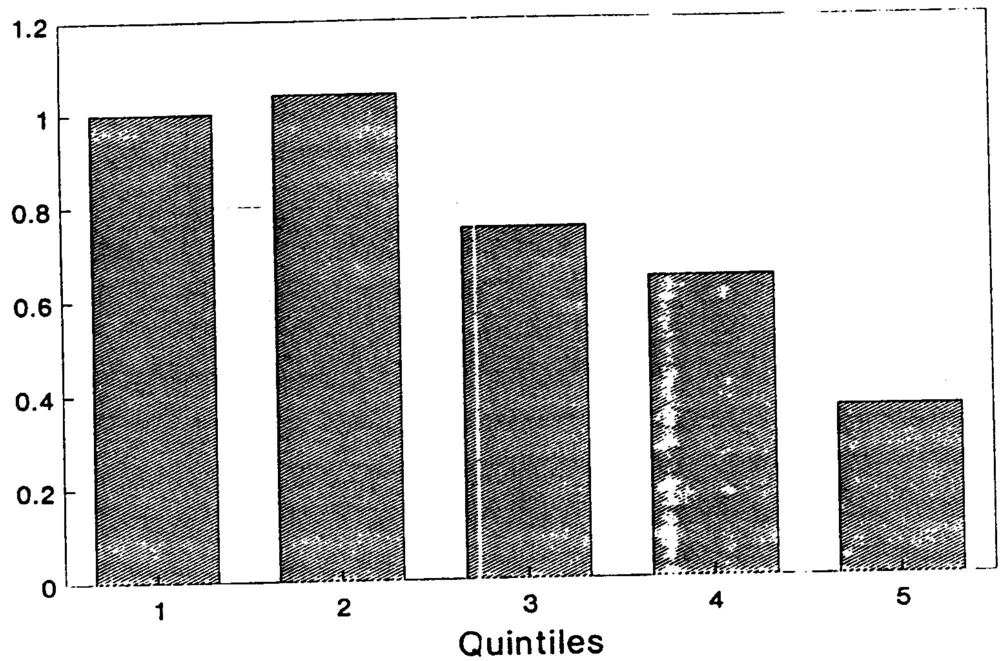
Rimm, EB, Giovannucci,E, Stampfer, M.J., Colditz, G.A., Litin, L., Willett, W.C. Reproducibility and validity of a self-administered semi-quantitative food frequency questionnaire among male health professionals. (submitted for publication)

Willett, W.C., Stampfer, M.J., Colditz, G.A., Rosner, B.A., Spiezer, F.E., Relation of meat, fat, and fiber intake to the risk of colon cancer in a prospective study among women. N Engl J Med 1990; 323:1664-1672.

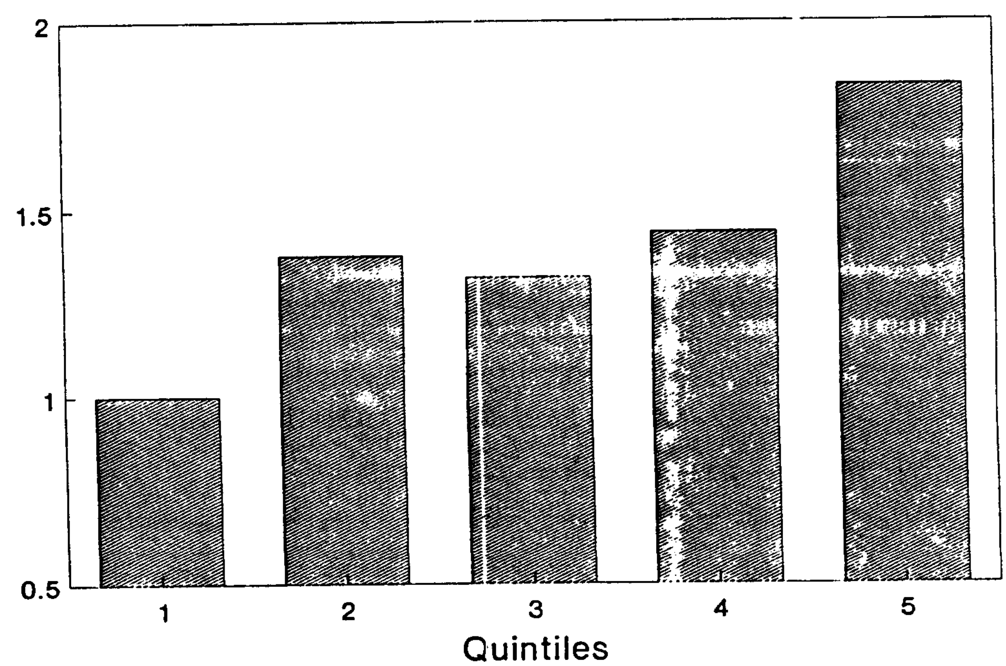
Giovannucci, E., Stampfer, M.J., Colditz, G.A., Rimm, E.B., Speizer, F.E., Willett, W.C. Relation of diet to the risk of colorectal adenoma in men.

# Dietary Fiber

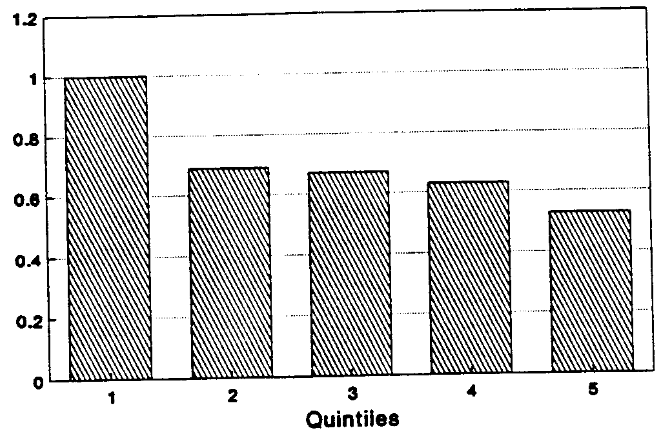



# Red Meat/Chicken & Fish

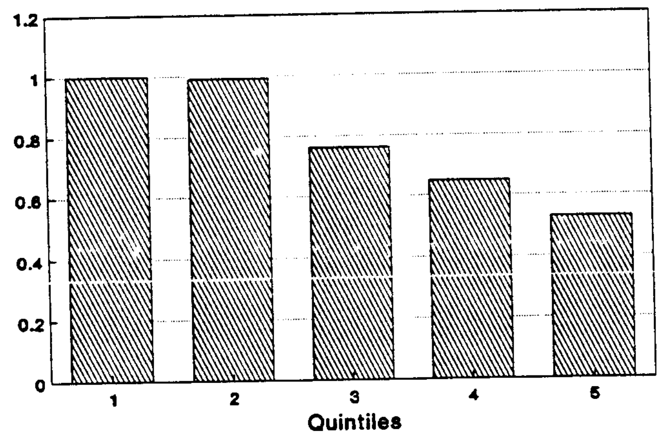



p value, test for trend = 0.02

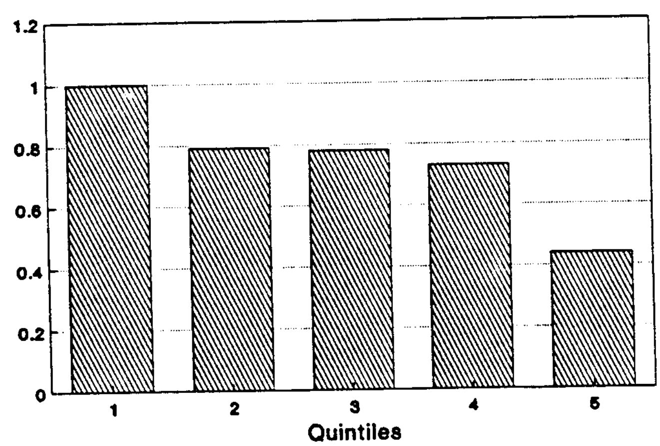
## Fruit Fiber




## Vegetable Fiber




## Cereal Fiber



# Current HNIS Nutrient Data Research

**Ruth H. Matthews**
**HNIS / USDA**
**Hyattsville, Maryland**

The primary focus of the Nutrient Data Research Branch of the Human Nutrition Information Service is maintaining quality and currency of data in the nutrient data bases including the Standard Reference Data Base, the Primary Data Set, and the Survey Data Bases. I want to describe several aspects of quality and currency maintenance in our program.

## Maintaining Quality of Data

*Percentage of Analytical Data*

One of the basic indices for quality of a data base is the percentage of analytical values. Hepburn[1] reported in 1987 on the percentage of analytical data in the Primary Data Set (PDS). The PDS contains nutrient profiles of basic food items taken from the USDA Nutrient Data Base for Standard Reference with additional nutrients and foods added from new analytical data and imputed with documentation.

The common perception is that analytical data are better in quality than imputed data.

The percentage of analytical data for proximates, calcium, iron, phosphorus, potassium, sodium, thiamin, riboflavin, and niacin is usually greater than 95 percent. Since the Hepburn report in 1987 over 900 items have been added to the PDS in response to requests from the Surveys. Some foods in the PDS contain low or negligible levels of certain nutrients (near level of detection for the nutrient) and contribute little to total food intake. However, other foods, such as milk and bread, contribute significantly to intake of several nutrients.

After food consumption data were obtained from recent surveys (CSFII 1985 and NFCS 1987-88) for each food item, total grams consumed were multiplied by the value of each nutrient.

The resulting figure provides the amount consumed for each PDS item. The foods that contribute to 80 percent of the total intake of each nutrient have been defined by NDRB as "key foods." These foods are important sources of nutrients of public health significance and are reported on here.

The percentage of analytical data increased considerably in the PDS for all nutrients between 1987 and 1991. In Table 1, you will note that the percentage of analytical data for vitamin B-6 and magnesium improved markedly to 92 and 90 percent from 72 percent for each of these nutrients in 1987. Copper, folate, dietary fiber, and vitamin E (alpha-tocopherol) remain less than 90 percent analytical. Alpha-tocopherol is lowest at 47 percent--up from 39 percent. Results of contract research (current and planned for next year) will increase the percentage of analytical data for alpha-tocopherol, making this nutrient comparable to the others.

*Research for Generating Data*

Since 1987, extramural research funds have been allocated for analyzing several species of farm-raised and wild fish, speciality fruits, miniature vegetables, and special cheeses; for strengthening data (increasing N) for mineral elements including selenium, copper,

**Table 1. - Percentage of Analytical Data from Primary Data Set**

| Nutrient | 1987[1] | 1991 |
|---|---|---|
| Vitamin C | 92 | 95 |
| Vitamin B-12 | 70 | 95 |
| Carotene | 88 | 93 |
| Vitamin A (RE) | 73 | 92 |
| Vitamin B-6 | 72 | 92 |
| Zinc | 79 | 91 |
| Magnesium | 72 | 90 |
| Copper | 71 | 88 |
| Folate | 69 | 87 |
| Dietary fiber | 40 | 82 |
| Alpha-tocopherol | 39 | 47 |

[1] Hepburn

manganese, magnesium, and zinc; and for developing and verifying data on vitamin E, dietary fiber, and nutrient retention.

Today, some species of fish are marketed primarily as farm-raised. Table 2 shows fat and cholesterol values for four species of fish from our recent contract. Fat is higher in these species when farm-raised. Cholesterol values are not significantly different.

**Table 2. - Fat and Cholesterol in Cooked Fish Wild vs Farmed**

| | Fat | | Cholesterol | |
|---|---|---|---|---|
| Species | Wild | Farmed | Wild | Farmed |
| | g/100 g | | mg/100 g | |
| Catfish | 2.8 | 10.0 | 72 | 64 |
| Trout | 5.8 | 7.2 | 69 | 68 |
| Salmon (coho) | 4.3 | 9.0 | 55 | 62 |
| Oysters | 1.9 | 2.3 | 49 | 38 |

*Research for Monitoring Data*

Monitoring contracts for key foods have been divided as shown:

1) Analyses of proximates, vitamins, minerals, dietary fiber.
2) Analyses of lipid components--including fat, fatty acids (including geometric isomers), and sterols (including cholesterol).

Foods to be monitored are selected according to the amount consumed, the number of nutrients where the food is among the top 80 percent, and other sources of available data. For some nutrients such as vitamin A, only a few foods are responsible for 80 percent of the nutrient. For others such as magnesium, an extensive list of foods provides 80 percent contribution. The top 10 foods for six nutrients are given in Tables 3 through 8, with the percentage of the nutrient intake provided by that food. Needless to say, these foods have the greatest influence on the total nutrient intake.

Foods such as potatoes in several forms, milk and enriched white bread, toasted enriched white bread, and rolls and buns are at or near the top of lists for several nutrients in percentage of total intake. Effects of enrichment are evident for thiamin and iron, showing enriched white bread and rolls and buns at or near the top. Whole milk is definitely a "key" food as shown by its contribution of 4.0 percent of the diet's vitamin B-6, 3.9 percent of the thiamin, 9.2 percent of the potassium, 5.5 percent of the zinc, and 2.3 percent of the sodium intake.

**Table 3. - Key Foods for Vitamin B-6 Intake**

| Food | Vitamin B-6 Intake % |
|---|---|
| Bananas | 4.4 |
| Whole milk | 4.0 |
| Potatoes, boiled w/o skin | 3.1 |
| 2% lowfat milk | 2.2 |
| Gr beef, reg, broiled | 2.0 |
| Eggs | 1.6 |
| Potatoes, boiled w/ skin | 1.4 |
| Baked potatoes | 1.4 |
| Cheerios | 1.3 |
| Total (wheat) | 1.3 |
| | 22.7 |

**Table 4. - Key Foods for Thiamin Intake**

| Food | Thiamin Intake % |
|---|---|
| White bread, enriched | 6.2 |
| Rolls and buns | 4.2 |
| Whole milk | 3.9 |
| White flour, all-purpose | 3.5 |
| White bread, enriched, toasted | 2.7 |
| Orange juice, frozen concentrate, diluted | 2.5 |
| 2% lowfat milk | 2.2 |
| Ham, cured, boneless, roasted | 1.7 |
| Ham, cured, boneless, unheated | 1.6 |
| Cracked wheat bread | 1.5 |
| | 30.0 |

**Table 5. - Key Foods for Iron Intake**

| Food | Iron Intake % |
|---|---|
| White bread, enriched | 4.4 |
| Rolls and buns, enriched | 3.3 |
| Kellogg Raisin Bran cereal | 2.9 |
| White flour, all-purpose, enriched | 2.7 |
| Beef, ground, regular cooked, broiled, medium | 2.2 |
| Eggs | 2.1 |
| Bread, white, enriched, toasted | 1.9 |
| Cheerios | 1.5 |
| Total (wheat) | 1.4 |
| Cracked wheat bread | 1.4 |
| | 23.8 |

**Table 6. - Key Foods for Potassium Intake**

| Food | Potassium Intake % |
|---|---|
| Whole milk | 9.2 |
| 2% lowfat milk | 5.0 |
| Coffee, brewed | 4.4 |
| Orange juice, frozen concentrate, diluted | 3.5 |
| Potatoes, boiled w/o skin | 2.4 |
| Bananas, raw | 2.0 |
| Potatoes, frozen, French fries | 1.8 |
| Potato chips | 1.7 |
| Tea, brewed | 1.4 |
| Beef, ground, cooked | 1.4 |
| | 32.8 |

**Table 7. - Key Foods for Zinc Intake**

| Food | Zinc Intake % |
|---|---|
| Beef, ground, cooked, broiled, medium | 5.7 |
| Whole milk | 5.5 |
| 2% lowfat milk | 3.0 |
| Beef, ground, lean, cooked, broiled, medium | 2.8 |
| Cheese, American, processed | 2.2 |
| Eggs, whole | 1.9 |
| Beef, ground, extra lean, cooked, broiled, medium | 1.8 |
| Beef, ground, regular, cooked, well done | 1.8 |
| Oyster, Eastern, raw | 1.2 |
| Ice cream, regular, 10% fat | 1.2 |
| | 27.1 |

**Table 8. - Key Foods for Sodium Intake**

| Food | Sodium Intake % |
|---|---|
| Table salt | 27.0 |
| Cheese, American, processed | 3.4 |
| White bread, enriched | 3.3 |
| Rolls and buns | 2.4 |
| Whole milk | 2.3 |
| Regular margarine, salted | 2.2 |
| Ham, cured, roasted | 1.5 |
| White bread, enriched, toasted | 1.4 |
| Tomato sauce | 1.3 |
| 2% lowfat milk | 1.3 |
| | 46.1 |

*Quality Assurance Program*

The Quality Assurance (QA) Program of the Branch includes four stages:

· A Three-Member Quality Assurance (QA) Panel coordinates and reviews work relative to QA including negotiating contracts for development of reference materials based on specific need. The Panel recommends QA steps during development of new contracts and serves on Technical Panels to review QA aspects of the proposal.

· Develop and Characterize Reference Materials

One member of the QA Panel serves as the Contracting Officer's Representative and negotiates contract for reference materials development and characterization. The three-member panel identifies appropriate materials to measure performance according to foods and nutrients being studied. At least one NIST Certified reference material is selected for each set of check samples.

· Stringent Screening of Contractors Before Award.

During evaluation of the proposal, a contractor's ability is proved through analysis of the check samples. Three reference materials must be analyzed for several important nutrients in a contract. Performance is measured against a standard reference value or range. Sometimes as few as 20 percent of prospective contractors are able to perform analyses as demonstrated by our screening program in the past 4 years.

· Annual Meeting of Contractors

Meetings are held each June to discuss problems in performance including sample preparation for analysis, plans for new work, sharing new developments and needs, or recommendations for new work. Special needs for problem solving and presentations on special subjects such as dietary fiber, vitamin A methods, and folate analyses are also part of the annual Contractor's Meeting.

*International Collaborative Study for Dietary Fiber Analyses*

In 1989, HNIS initiated an international collaborative study to ascertain what methods of analysis may be used for dietary fiber compilation. The methods compared were AOAC, Englyst, Mongeau, and simplified AOAC (NCL). The AOAC method of dietary fiber analysis was conducted in several laboratories by analyzing 25 U.S. foods in blind duplicate. The food samples were prepared in one lab and sent to all participating labs. Foods selected included some that had already presented problems in analysis. High-fat, high-starch components presented some problems for some laboratories. Participants were Eric Florence in Reading, England; Roger Mongeau, Canada; J. Robertson, Cornell University, D. Gordon, University of Missouri; J. Augustin, University of Idaho, and Betty Li, Nutrition Composition Laboratory, Beltsville, MD. All laboratories were provided with the same lot of Sigma kits (enzyme). I reported on these results at FASEB in 1990. Other methods were used for analyses of dietary fiber on the same food samples. Hans Englyst (Cambridge, England), Roger Mongeau and Dennis Gordon analyzed by the Englyst method; the latter two by the Mongeau method; and Augustin and Li, by the simplified AOAC procedure. Results showed good agreement for most labs using the AOAC procedure.

*Consultant Panel*

Early in 1991, HNIS appointed a three-member Consultant Panel for Nutrient Data Research. They represent industry, academia, and government. Their areas of expertise are data base management, data base building, and nutrient analytical methodology skills.

*The Consultant Panel's role is to:*

- identify issues of critical importance to maintain and enhance nutrient composition data
- and address issues NDRB would like recommendations on

*The current members are:*

Norman Bednarcyk, Nabisco Brands

Loretta Hoover, University of Missouri

A.C. Soliman, FDA, Atlanta, Georgia

**Maintaining Currency of Data**

*Annual Supplements to AH-8*

The first (1989) Annual Supplement to AH-8 was issued in 1990 to update values in sections already published and to add new foods and new data. In the first Supplement, calcium and manganese tables of values were added for AH-8-1, Dairy and Egg Products, and for AH-8-2, Spices and Herbs. We often receive questions about the date of publication, i.e. (-AH-8-1, Dairy and Eggs, published in 1976). Supplements are issued each year to keep data current. So far, 210 new or revised food items have been added including revised data on eggs as well as tables of dietary fiber and new guides. The second supplement, just published, has 80 new foods including 46 fish items and 34 revised foods including new data for light tuna in water and selected fruits and vegetables. Some new foods that have appeared in the Supplements are goat cheeses, miniature vegetables, asian pears, and new vegetable and fish oils.

The third Supplement will include approximately 110 fresh pork cuts based on recent research on market samples. It will also include cornish game hen, farm-raised fish, and other items.

*Revised Sections of AH-8 and Regular Database Update.*

The AH-8-13 on Beef Products was completely revised and published in June 1990. The data were available earlier that month on our Nutrient Data Bank Bulletin Board. It is operated as a public service to provide information about current HNIS publications and computer files on the nutrient composition of foods, as well as announcements about this conference and other relevant topics.

AH-8-8, Breakfast Cereals and AH-8-4, Fats and Oils revisions are planned for the future and willinclude new items and revised data. When new sections or supplements are issued, the data bases, are updated in new releases.

*Other Publications*

Next year we plan to issue an abridged AH-8 and a Master Index for ease in locating the food items in the data bases.

*Industry Cooperation*

NDRB has maintained close working relationships with industry over the years. All sections of Handbook 8 have been reviewed by industry representatives to assure no changes in production or formulations. For example, over 30 companies reviewed AH-8-19, Snacks and Sweets, in addition to trade associations such as, the Chocolate Manufacturers Association and the Snack Food Association. Over 30 industry groups are reviewing the Baked Products section, AH-8-18.

Some studies are planned with industry so that generated data will be suitable for our needs. We have cooperated with Texas A&M University, the National Livestock and Meat Board, United Egg Producers, Snack Food Association, and the Produce Marketing Association to discuss sampling and handling of the materials. For analyses of meats, we developed and recommended a standard protocol for nutrient analyses which has since been adopted by the American Meat Institute. The protocol calls for standard methods for storage, dissection, and handling of samples before, during, and in preparation for nutrient analyses.

**Future Plans**

-- Continue to monitor key foods -

Contracts on proximates, vitamins, minerals, and dietary fiber

Contracts for lipid components

-- Strengthen special data bases on:

Vitamin E--present total vitamin E as alpha-tocopherol equivalents and individual tocopherols and tocotrienols as needed.

Dietary fiber--including soluble and total dietary fiber

Vitamin K

-- Build a data base on fat-substituted or fat-replaced foods

Baked products

Salad dressings

Frozen desserts

Spreads

Other dairy items

Top brands of these foods varying in fat replacement or fat substitution components will be analyzed. Components include Simplesse , medium chain triglycerides (MCT), potato starch, rice starch, maltodextrin, polydextrose, carrageenan, guar gum and locust bean gum, and Olestra when approved by FDA.

As we move into the 21st century, effects of advances in biotechnology will become more and more evident as well as other technological advances in food production and processing which undoubtedly will markedly influence nutrient composition of foods. USDA will continue to monitor these advances and keep nutrient compostion data bases current and accurate.

[1] Hepburn, F.N. Food Consumption/Composition Interrelationships. Adm. Rpt. No. 382, pp 68-74. 1987.

# Management of Food Composition Databases in Foodservice Settings

**Loretta W. Hoover, Ph.D., R.D.**
**University of Missouri-Columbia**
**Columbia, Missouri**

Food composition databases serve important functions in foodservices. If the operation supports patient foodservice in a healthcare setting, clinical staff need estimates of the nutritional content of the foods served to their patients as they plan modified diet menus or tailor menus for a particular patient. In other foodservice settings, nutrition-conscious clients are requesting information about the nutrient composition of menu offerings. Often the food products offered are mixed dishes prepared on-site from a variety of ingredients. Some of the ingredients may be formulations developed specifically for quantity food production operations and not included in standard food composition databases. Other ingredients may be foods in an unprocessed form requiring pre-preparation and estimation of weight changes prior to incorporation in mixed dishes. Foodservices are not likely to have the nutrient composition of recipes analyzed in a chemical laboratory; estimation of nutrients in recipes with a computerized database system is the most common method for determining nutrient profiles. Thus, access to a food composition database facilitates operations in the foodservice setting.

Management of a food composition database in the foodservice setting involves some of the same considerations encountered in other settings. Although not concerned with preservation of previous editions of the database, the database manager in a foodservice setting is concerned about maintaining an up-to-date database that reflects ingredients and ready-to-serve foods used in the operation. The updating activity requires effective quality control procedures and well-designed software to accomplish the maintenance functions. In organizations providing both foodservice and patient/client care, integrated applications eliminate the expensive and time-consuming tasks of maintaining multiple data bases. These considerations are addressed more thoroughly in the following sections.

*Establish a method for identifying added nutrient profiles*

Multiple sources of data are often maintained for the same food item in a food composition database and require methods for identifying and organizing supplemental data. If additions are necessary to identify a particular ingredient used in the foodservice operation, some system is needed to position the profile for that item in the database so that it will not be lost when the standard items in the database are updated. Numbering systems or qualifiers are techniques for protecting these additions in a database over time. When the new data are for the brand name version of a generic item listed in the food composition database, a source qualifier could be coded as a part of the data record to distinguish that food item from the one in the standard source database (eg. 06253 00 for a USDA record for the generic food item; 06253 01 for an added record for brand specific data ). When a corresponding generic item is not present in the food composition database, one is challenged with positioning the new profile in the database so that subsequent updates from the standard source will not result in "over-writing" of that data record. In positioning the new data record, some system is desirable so that new items are in logical locations in the database or on listings of the nutrient database (NDB).

*Integrate applications so only one NDB is required*

Maintenance of a food composition database can be centralized in an organization if applications are integrated. Thus, the duplicative efforts of database maintenance can be avoided. However, if both foodservice and other uses will be made, the requirements of all applications must be considered when the database structure is designated and the software applications are developed. With careful system design, a foodservice operation can utilize the same food composition database used to support research or patient care functions. Centralized management of the updating process with strict quality control procedures will probably result in a more reliable and economical source of data for all aspects of an organization.

*Coordinate NDB with ingredients in a recipe data base*

Food production practices influence what nutrient profiles will be required in a food composition database. Also, the nutrient calculation method will influence what food composition data are used to estimate nutrient profiles for a mixed dish. Frequently foods are served with non-edible parts (eg. bone in a pork chop); the non-edible portion needs to be included in the portion weight in order to monitor recipe yield and portion control. However, the nutrient values for the portion should correspond to only the edible portion. Thus, for foodservice operations, intermediate and finished recipe and portion weights are useful.

*Develop output or applications to support menu planning and/or menu tailoring*

Nutrient databases and recipe databases provide the infrastructure in a foodservice system and support menu management functions. Once the databases are loaded with data that accurately reflect the foods served in an organization, software applications can be developed to facilitate both strategic and tactical aspects of foodservice planning and service. Cross-references of ingredients and recipes can be useful to identify recipes containing specific ingredients. On-screen analysis of menus can be used to determine if a specific combination of foods contain nutrients at specified levels. The applications can extend the functionality of the database system.

*Use yield factors to reflect cooking/preparation loss and facilitate purchasing activities*

Yield factors are especially important in a database system supporting food production activities. Foods may be purchased in bulk in an unprocessed form but stated according to some processed amount in a recipe. In these instances, a yield factor is necessary in the database system to convert the amount stated in the recipe to the amount that must be purchased for the recipe. This yield factor is not involved in computing the nutrient profile for the recipe but is required to correctly cost the ingredient in the recipe when some pre-preparation occurs in the foodservice facility.

# Maintenance of Food Composition Databases

**Sally Schakel**
**Brian Westrich**
**University of Minnesota**
**Minneapolis, Minnesota**

**Introduction**

The Nutrition Coordinating Center (NCC) at the University of Minnesota maintains two food composition databases: 1) a Nutrient Database that contains descriptions, nutrient values and reference codes for representative foods in the American diet, and 2) a Brand Name Database that contains nutrient, ingredient, density and serving information provided by food manufacturers for brand name products. The Brand Name Database is used to provide information for foods in the Nutrient Database and to aid in determining coding for brand name products by matching them to similar Nutrient Database entries.

**Maintenance of the Nutrient Database**

The NCC Nutrient Database contains approximately 1500 elemental (non-recipe) foods and 1000 recipes or formulations of commercial products. For more efficient maintenance of the Nutrient Database, entries are limited to foods in the form in which they are consumed (e.g. entries include only cooked, not raw potatoes), and foods with similar nutrient contents are grouped together into a single database entry. Components of each database entry include a food code, food description, nutrient values per 100g of food, and reference codes for each nutrient value.

The Nutrient Database is used to calculate dietary intakes for numerous research studies and must be maintained to meet the following needs of these studies:

1) *Provide nutrients of interest.*
The Nutrient Database was developed in 1974 to calculate dietary intakes for two major cardiovascular studies. Because these studies were especially interested in intake of total fat, fatty acids and cholesterol, the database reflected that emphasis by including these nutrients along with the other proximates and selected vitamins and minerals. Since that time, requests have come from research studies dealing with hypertension, cancer and diabetes to use the Nutrient Database. To accommodate these studies, additional nutrient fields, such as sodium, dietary fiber and sugars, have been added. Currently the database contains 93 separate nutrients, including 23 individual fatty acids and 18 amino acids.

2) *Provide specificity of foods to obtain differences in nutrients of interest.*
When a new nutrient is added to the database, existing database entries are often split into two or more entries so that differences in the new nutrient of interest can be seen in otherwise similar foods. For example, the database was expanded to include separate entries for canned and frozen vegetables when sodium became a nutrient of interest. When dietary fiber was added to the database, the number of cereal entries increased to account for different levels of dietary fiber in these products, and entries for pasta expanded to white and whole wheat.

3) *Provide an updated database.*
As new or better nutrient values become available, it is important to update the Nutrient Database with the more current data. The primary source of data is the USDA Nutrient Database for Standard Reference. Each NCC food that is based on a USDA entry contains that USDA code number in the field for "reference source." Values from a new release of the USDA database can be transferred directly into the NCC database using the reference code as a link. Differences between the old and new USDA values are flagged and verified as correct before the new value becomes a part of the NCC database. Additional data sources used are other USDA publications such as handbooks, provisional tables, and the survey database; scientific literature; foreign food tables; and food manufacturers via the Brand Name Database.

4) *Provide a complete database.*
Because missing values are calculated as zeros in dietary intakes, it is important for each database entry to have a complete nutrient profile. It is better to have a reasonable estimate of a nutrient amount for the food than leave the field blank. This often involves imputing nutrient values. Missing values can be imputed from a similar food, from a different form of the same food (e.g. raw to cooked using retention factors), from a related nutrient, or by developing a formula of the product ingredients that can be used to calculate the missing nutrients. For example, a manufacturer may provide basic nutrient values for a cocoa powder, but no values for sugars, dietary fiber or fatty acids. The ingredients of the cocoa powder are entered into a computer program in the order that they appear on the label. Using food formulary books or a reasonable guess, the amounts of

each ingredient per 100g are entered. Then the nutrients given by the manufacturer are entered so that the total nutrients of the ingredients can be compared with the nutrients given by the manufacturer. Guidelines for an acceptable difference between a calculated nutrient value and that of the manufacturer have been established. If the nutrient comparisons are not close enough, the outlying nutrient is flagged by the computer and the formula must be adjusted until all nutrient differences fall within acceptable limits. When the formula has been determined, the computer calculates all missing nutrient values by totaling the values from the ingredients in the formula. These calculated values are then entered into the *Nutrient Database* entry. Each nutrient value is accompanied by a source code so that imputed nutrient values can be recognized and replaced with analytic data when they become available.

5) *Provide an accurate database.*
Validation of the Nutrient Database involves several procedures used to locate errors. First, edit limits are used to identify values incorrectly entered into the database. Within the computer program are maximum allowed nutrient values for foods within a particular food group. These limit values cause the computer to flag any nutrient value that appears too large for the type of food that is being entered. The nutritionist must check all flagged nutrients and verify that the values are correct.

The most thorough validation check is a review of a new entry or changes to an existing entry by a second nutritionist. All values for nutrients, serving sizes, and densities are checked against the original source of the data. Calculations and assumptions made about the food are corroborated by this nutritionist.

A third series of checks are run on the internal consistency of the database before a version of the Nutrient Database is released. Calculations are done to compare the total weight of the proximate nutrients to 100g; to compare actual calories with calculated caloric values; to compare weight of individual fatty acids to total fat; amino acids to protein; and fiber, sugars and starch to total carbohydrate. Large differences are flagged by the computer and the nutrient values for that food are verified by a nutritionist. Other nutrients, such as vitamins and minerals, are checked by comparing values of foods within the same food group (e.g. compare riboflavin of all dairy products) and identifying outliers that need to be checked.

**Maintenance of the Brand Name Database**

The first step in maintaining a brand name database is to contact manufacturers to acquire brand specific nutrient data, a process detailed in an earlier presentation. NCC also refers to printed publications and contacts food retailers to obtain brand name product information. Data are obtained for over 6,000 brand name products annually.

Because more than 12,000 new brand name products are introduced into the market each year, the next step in the process is to determine which of these products need to be entered into the database. NCC has developed criteria to help make these decisions. For example, the Brand Name Database includes only foods whose nutrient content varies significantly from brand to brand. Therefore, different brands of cookies are included in the database, while different brands of canned peaches are not. In addition, only nationally distributed products are included in the database.

The Brand Name Database is designed to a) allow entry of data in the same format as it is supplied to NCC, b) preserve time-related data, and c) provide data validation at the time of data entry. Nutrient values received from manufacturers are often expressed using different units of measure (grams vs. milligrams) and methods of reporting (nutrients per serving vs. nutrients per %USRDA). The Brand Name Database allows entry of all types of values. In addition, multiple values for each nutrient can be stored in the database. For example, both analytical and label values for the same product can be stored, as well as nutrients both per 100g and per serving size. During data entry, the source of the nutrient information can be specified by selecting from over 50 different sources, including product label, calculated, or analytic data. Because the Brand Name Database can store an unlimited number of nutrients for each product, and each nutrient can have an unlimited number of values, a date is entered with each value. This allows use of the most up-to-date nutrient information available, while at the same time preserving an historical record of changes in foods over time. A "reason" code notes whether newer information is due to better data or to a change in product formulation. Currently, only limited data validation occurs at the time of data entry. This includes checking the validity of codes entered for nutrient name, source, and method of measurement. However, in the near future, a more thorough system of quality control will be implemented similar to validation procedures used for the Nutrient Database. In addition, quality control checksspecifically designed to handle the multiple nutrient values possible in the Brand Name Database will be added, as well as duplicate entry of randomly selected products.

Sometimes manufacturers do not provide sufficient data to calculate nutrient values. For example, nutrient content of a candy bar may be reported per bar, but the weight of the bar is not provided. In the case of missing data, NCC contacts the food manufacturer for this additional information or obtains it from the product label. Currently, the Brand Name Database contains over 5,000 brand name products, 127 different nutrients, and 70,000 individual nutrient values. The database size is projected to expand to 6,000 products and 90,000 nutrient values by the end of this year.

Reports can be generated from the Brand Name Database to match brand name products to the most appropriate entry in the Nutrient Database. This matching is based on nutrient composition. For each food type, "key" nutrients are selected for which the food type is a significant source. For example, calcium is a key nutrient for dairy products. Key nutrient values for each brand name product are then compared with the corresponding nutrient values for existing entries in the Nutrient Database. Differences must be within an acceptable range established for each nutrient. If there is no match between a product and a Nutrient Database entry, a new Nutrient Database entry may be created. This new entry is usually based on a product formula developed using the procedure described in section 4 above. These practices permit NCC to accommodate nutrient differences between brand name products without greatly increasing the number of Nutrient Database entries.

# Management of Food Composition Data Bases

**Brucy C. Gray**
**HNIS / USDA**
**Hyattsville, Maryland**

The first item of consideration in the management of a database is the use to which the database is to be put. A database in a retail store may be used to obtain the price of an item as it is sold and to maintain an inventory of the stock on hand. Except for the stock on hand and price, the actual values of fields in the database remain unchanged throughout its existence. Any new item in that product line would be inserted as a new item in the inventory although it may be a slight modification of an existing item. In this usage, the user system needs the capability of reading all fields in the file and updating the inventory count field and the cost field. The user system, along with its other tasks, needs an efficient way of adding new items to the file and locating a particular record in the file.

Another type of data base is one which is quite dynamic, needing to be updated on a continuous basis. Such a database may be used by a researcher or a medical doctor who maintains an automated record of his patients and the history of an ailment along with the effect of a treatment that is being provided. Such a database as this needs the capability of being constantly updated with expanded field sizes and the capability of comparing entries for different records for the same individual along with making comparisons across individuals. The need for instant access to a particular record is not as great as that needed for the retail store sales but is greater than what is provided by sequential reading. For the convenience of making comparisons, it may be necessary to translate the data being entered into predefined special codes. There is a need to be able to enter text fields as well as a capability of pulling up a specific text field for review at the request of the user. There may be a need to retrieve all records from the data base for the cases where predefined fields contain specific codes or responses.

The message in the above two examples is that there is no single database or response that meets all of the needs of the users. Each application needs to be analyzed and a system developed which best meets those needs. This also implies that the management of the database is correlated with the way the database is set up and the uses to be made of that database. The remaining portion of this presentation will deal with the specific applications of the food composition databases that we have in HNIS.

The most popular of the databases that we distribute is the USDA Nutrient Data Base for Standard Reference. This is the machine readable form of the data contained in the Agriculture Handbook Number 8 series. The database itself is created from files produced in the Agency's Nutrient Data Bank system. That system has been discussed many times and I will not go into details about it at this time. Yesterday afternoon Betty Perloff gave an update to the status of that data bank.

Out of that databank and the system for producing the distributed copy, we receive the necessary files for the data and for the coding manuals. An automated system exists for producing files in the format for distribution. This is a rather straight forward system requiring only that the master file be retained in a machine readable form. The variable form of the system resides in the databank itself.

Two other composition data bases that we distribute for public use are the USDA Nutrient Data Base for Individual Food Intake Surveys and the USDA Nutrient Data Base for Household Food Use Surveys. The system of maintenance for these two are quite similar. The procedure for creating and maintaining the machine readable form of these data bases is automated. Its operational principle is based on the ability to determine the nutrient content of foods by starting with the nutrient content of the raw ingredients and obtaining the nutrient content of the finished product by appropriate adjustments to account for the gains and losses that occur in the preparation process. The automated form of this system consists of files of all foods that may be used as ingredients in recipes, and the developed automated procedure for processing the files.

At the start of a survey, files exist that contained all of the foods anticipated to be consumed by participants in the survey. For all raw ingredients, the file contains the nutrient content of the ingredient. For foods that are not consumed as raw ingredients a recipe file exists which contains the ingredients used in the prepared food, the respective quantity used, the appropriate refuse factor, the nutrient retention code and other information that impacts on the final nutrient level in the consumed food. The system is structured around performing the needed computations and producing a finished record of the nutrient content of the eatable product. To use this recipe file a system has been designed which takes the ingredients in each recipe and calculates the nutrients in the finished product after adjusting for moisture gain or loss, nutrient loss due to heat and the absorption or loss of fat or other substances.

In this system, the format of the files is fixed. There is a need for immediate access to the records of the ingredients, the supporting files of the nutrient retentions and the refuse factors. There is also a requirement to access the files of the system as needed. This system must also have the capability of being updated as needed with the capability of using the result of previously generated recipes and ingredients of new recipes. This system differs from the two mentioned above in the sense that a part of its structure is a well defined computational algorithm. It takes in information in a fixed format and, as needed, puts out information in a fixed format. The system does not have the capability of expanding the field sizes in order to hold more information and the retrieval of information is not based on the description of the food.

In summary, databases differ based on their content, their size, as well as the intended use. Because of the sizes of the files and the specialized computations that are required, ours have been developed to permit easy handling. It permits them to be stored in an inexpensive form and gives the desired control for maintenance. This has been done by writing specialized programs to perform the handling. Applications centered around the use of Database Management Packages can usually do all that is needed. The drawback on that is that the applications systems have to be developed and the cost of maintenance is greater than the cost of maintaining sequential files. The guidance that I would give is to discuss the files and the manner in which they will be used with a person experienced in computer operations and let that person provide guidance on a system suited for the applications intended.

# Using International Nutrient Data An Overview

**Suzanne P. Murphy**
**Doris H. Calloway**
**University of California, Berkeley**
**Berkeley, California**

Over the last few years we have worked on several projects in which we needed nutrient data for foods from countries other than the U.S. This overview will discuss some of these experiences. Of necessity, our view will be that of a U.S. researcher, but we believe our perspective would apply to researchers in other countries as well.

In general, nutrition researchers need international nutrient data in two situations. Obviously, when conducting research on populations residing outside one's home country, it will be necessary to locate nutrient data on foods of the study country. We have been involved in the Nutrition Collaborative Research Support Program in Egypt, Kenya, and Mexico[1], and together with other U.S. investigators, as well as scientists from the host countries, have spent considerable effort in locating appropriate food composition data for the rural villages being studied. We have been aided by several published food composition tables from these three countries, but in some instances, there were significant problems in correctly identifying the local food items and estimating their nutrient content.

More to the point for this discussion, nutrition researchers may need international nutrient data even when conducting studies within their home countries. This occurs when foods from other countries appear in the diets of the home country population. In the U.S., this situation occurs commonly; imported food items are commonplace in virtually all food stores, and in most people's diets. The variety is astounding, including such diverse products as canned nopales from Mexico, pesto sauce from Italy, and salted plums from Asia. Many of these imported foods have no parallel in the U.S. food composition tables.

Another changing aspect of the American diet, and probably of those of most other countries as well, is the increasing popularity of ethnic restaurants. Here the problem of locating food composition data is even more difficult, because researchers must worry not only about the nutrient content of the ingredients, but also the proportions of the ingredients in the final dish consumed by the patron. Often a subject trying to fill out a dietary record or recall will know neither the ingredients nor the proportions. Thus, we have a need for information about the usual content of mixed dishes from other countries.

Lastly, we have occasion to collect dietary data for people who have emigrated from another country. These people may grow foods from their native countries, and may prepare these foods, as well as local foods, in ways not customarily considered by U.S. food composition tables. Sometimes, even food composition data from the country of origin may not be useful because dishes are modified to include a mix of locally-available foods as well as foods from the native country. For example, American varieties of chiles may be substituted for native Mexican chiles in various mixed dishes such as enchiladas or mole sauce, so the vitamin content listed in a Mexican food table for these mixed dishes may be inappropriate for an Americanized version of the recipe.

Nutrient data from other countries are usually obtained from published or printed tables. INFOODS has published a directory of food tables by country which I have found very useful[2]. Sometimes important data may be found in research articles in journals; often these articles address the content of selected foods or nutrients, but they still may be very useful when compiling food composition data. Occasionally nutrient data are obtained, by request, from other investigators who have not had occasion to publish their food composition data. We believe that exchanging nutrient data (whether published or not) via electronic media will become more and more common as we all have better access to computer networks, and we hope this mode of data exchange will be greatly expanded in the future.

There are several types of information that ideally would accompany all nutrient databases. First, when information is exchanged between countries, the food descriptions are crucial. How to accurately describe foods has been an item of considerable discussion, as you have heard from other speakers at this conference and others. However, it seems that a minimum food description would include the name in English (speaking as American investigators) as well as the local name, plus the biologic name and brand name if a processed food item. Maturity may be important for some foods (e.g., mature vs. immature beans), as well as the part consumed (leaves, roots, stems, etc.). One would need to know any processing or preparation the food has received as well as any fortification or enrichment (we have found this information is often difficult to determine from local descriptions). Also, we would suggest that a description of typical recipes is important for mixed dishes that may be included in the table. Even for items such as bread, a recipe may be very useful in determining if the bread is whole grain or refined, what type of leavening agent (if any) is added, if fat is added, etc. Finally, it is important to know if the nutrient data for a food item represents a national average of that food from different regions, different seasons, etc., or if it is only representative of a local food item.

Food composition tables from other countries are not useful unless the nutrients are adequately descibed. Obviously, it is crucial to know the units for the nutrients, as well as any conversion factors that may have been applied. For example, there is considerable confusion today when using vitamin A values from international tables. The analytic methodology has changed over the last few years, so data from 20 years ago may not be valid. Furthermore, opinion on the appropriate availability factors to use with the various carotenoids has changed; we now use 6:1 for beta-carotene, whereas a few years ago we assumed 2:1. If a table combines retinol and carotenoids into a single "vitamin A" value, the results may vary widely depending on the conversion factor used.

Other desirable descriptions include information about the actual nutrient value given in the table. If it is a mean of multiple samples (the usual situation), then information about the range, the standard deviation, the median, etc., will provide useful information to other users. The number of samples should be given as well. The analytic method used should be described (by food item, if different methods are used for different foods). Finally, it is helpful to give the date of analysis (so the user can decide if the values are likely to be out of date).

These are lists of desirable documentation; probably no database has all the features described above. However, even if not available in electronic form, access to the information at least in printed form should ideally be available.

In closing, we would suggest some topics for discussion during the rest of this session on international data uses. (1) What types of international data are available at this time? (2) How would investigators from other countries obtain them? (3) What problems might investigators outside the country of origin encounter when using these data and what are the solutions to those problems?

### References

[1]Calloway DH, Murphy SP, Beaton GH. Food intake and human function: A cross-project perspective of the Collaborative Research Support Program in Egypt, Kenya and Mexico. University of California, Berkeley. 1988.

[2]Rand WM. International directory of food composition tables, 2nd edition. INFOODS, Cambridge MA. 1988.

# Using International Nutrient Data

**Barbara Burlingame**
**OCEANIAFOODS**
**Palmerston North, New Zealand**

This talk will introduce OCEANIAFOODS, and identify nutrient data products available from countries within OCEANIAFOODS and in what form they may be obtained. More importantly, this presentation will discuss the problems associated with international interchange of data, and some of the possible solutions to these problems.

## INTRODUCTION

OCEANIAFOODS regional group of INFOODS includes the countries Australia and New Zealand; and Papua New Guinea, Fiji, and the smaller Pacific Island countries represented by the South Pacific Commission in New Caledonia.

Australia is represented by a few major groups. The Department of Community Services and Health is responsible for developing the national nutritional database and its associated computer and printed products. The Australian Government Analytical Laboratory is responsible for much of the nutrient analyses work, along with researchers from the University of New South Wales and other organisations.

New Zealand is represented by the Department of Scientific and Industrial Research, which develops the national food composition databases, and conducts and coordinates the nutrient analyses work. Analyses are also conducted by the NZ Dairy Research Institute, Meat Industry Research Institute of New Zealand, and a few universities.

The South Pacific Commission operates the Pacific Island Food Composition Programme. Development of the nutrient database for the South Pacific and production of printed materials takes place at the Commission site. *In Fiji, the analytical work is conducted at the Institute of Natural Resources, University of the South Pacific; and in Papua New Guinea,* analyses are conducted at the Department of Agriculture and Livestock, Port Moresby and the National Analytical Laboratory, University of Technology, Lae.

## FOOD COMPOSITION DATA PRODUCTS

### *From Australia*

Products currently available include the printed *Composition of Food, Australia*, in five volumes to date (1-5); NUTTAB series, data files (6); and the recently released book *Nutritional Values of Australian Foods* (7).

### *From New Zealand*

Products currently available in New Zealand include the Food Composition Database as compressed datafiles on disk (8), datafiles with simple application software (9) and more advanced software (10), printed unabridged *New Zealand Food Composition Tables* (11), and users' guides for all products. There are three volumes in the *Composition of New Zealand Foods* series of books completed (12-14), with another several pending. Additionally, there is a consumer-style book with data on seven nutrients (15), and a book on the proximate composition of nearly 100 NZ fish species (16).

### *From the South Pacific*

Products available from the South Pacific Commission include newsletters, leaflets and information circulars on Island food and its composition. It is hoped that in the near future, the Pacific Island Food Composition Database will be available. This work is funded by international aid agencies, and has been continually cut in favor of other health projects such as Aids prevention.

## PROBLEMS WITH INTERNATIONAL INTERCHANGE

Within our own countries we have only the typical *problems experienced by all database developers.* When we engage in intra-regional interchange we need to *address other problems, and when contemplating in*ternational interchange we are faced with still more problems.

*The types of problems*

1. Naming of foods has long been acknowledged as a source of difficulty in development and use of extra-national food composition tables. Language is a problem, but even in the English language there are many differences between British English and American English. Australia and New Zealand would lean more toward British English, but still have unique descriptors for foods which can be confounding for other users. For example, the food record "New Zealand pumpkin" is not at all what is known to North Americans as pumpkin. The problem of naming or describing foods is being addressed by international bodies.

2. Naming of nutrients is potentially a more serious problem. A simple comparison between the USDA, the Australian, and the New Zealand tables shows some confusing anomalies. For example New Zealand lists a food component as "Available carbohydrate" and Australia and USDA list a component as "Carbohydrate, total". Although named differently, the Australian and the New Zealand (both of which exclude dietary fibre) are more similar than the identically-named Australian and USDA (which is calculated by difference and includes dietary fibre). This issue has been addressed by INFOODS in the book *Identification of Food Components for INFOODS Data Interchange* (17). In a datafile, each of these components would receive a different identifier: USDA's would be <CHOCDF>, which is "carbohydrate, total; calculated by difference"; Australia's would be <CHOAVL>, which is "carbohydrate, available"; and New Zealand's would be <CHOAVLM>, which is "carbohydrate, available; expressed in monosaccharide equivalents". The values all represent different things, and without the use of specific identifiers, international interchange will lead to dramatic misinterpretation of data by users in different countries.

3. Method of analysis is another problem area, which is partially addressed in the naming of nutrients. The most discussed component for which different methods yield very different values is dietary fibre. The above-mentioned INFOODS book also addresses this by listing eight different tagnames for fibre based on method, including method "unknown". This is essential for interchange, but also useful within a country where data presented as dietary fibre have been determined by more than one method.

4. Units used can also present a problem. When users are familiar with their own country's units for nutrients it is easy to miss a difference when using data from other countries. For example, manganese is expressed in milligrams by Australia and in micrograms by New Zealand. The Germans will often express the same nutrient (eg, sodium) in both grams and milligrams, depending on the amount present. It is a simple matter to catch differences in printed food tables, but tedious to try and determine differences in datafiles. Again, this problem area is addressed in at least two INFOODS' books published by United Nations University (17, 18), and all INFOODS-recommended tagnames for food components are unit- specific.

5. Copyrights

Most of the world's food composition databases and printed food tables are copyrighted. New Zealand reproduces some of the British and Australian data with permission from the copyright holders. Royalty payments and exchanges are involved. Software developers and book publishers using the New Zealand source data enter into arrangements with the Department of Scientific and Industrial Research for reproduction of these data. International interchange arrangements must include restrictions on reproduction of information, and permission when this is desired.

6. Uniqueness

Many foods are unique to a country or at least to the food database of a country with no equivalent counterpart (eg, New Zealand lists two varieties of feijoa, a fruit). Additionally, conditions affecting nutrient composition can also be unique. New Zealand, for example, is known to have unusual geochemistry, affecting the elemental concentration of foods. New Zealand also has unique food legislation which affects composition by regulating the extraction rate for refined grains, prohibiting nutrient fortification and enrichment of most foods including milk and refined grain products, and setting a minimum fat content for milk products. This poses no problem when viewing or comparing compositional data from different countries, but it would pose problems if data were to be adopted for use in another country's national database.

## SOLUTIONS TO THE PROBLEMS

The first step in solving the problems of international interchange of nutrient data is the adoption of international nutrient tagnames as defined by INFOODS (17). This eliminates the potential problems associated with inconsistencies in nutrient names, methods of analysis, and units.

The second most important step is the adoption of a standardized data format, which has also been defined by INFOODS (18). This would allow a programmed structure to be created and used by all involved in interchange. Data from around the world could simply be dropped in for easy and systematic storage, retrieval and manipulation.

Because most of the proposed solutions are untested to date, an interchange trial should be established. This should be undertaken by two countries from different regions, under the auspices of INFOODS and the United Nations University.

The successful trialing of the system (perhaps including the food descriptor system, LANGUAL) would lead to general adoption and dramatic easing of the problems now facing those who wish to participate in international interchange of food composition data.

## REFERENCES


1. Composition of foods, Australia (1989). AGPS, Canberra. Cat. No. 88 1488 5.

2. Composition of foods, Australia. Volume 2 Cereal and cereal products (1990). AGPS Canberra. Cat. No. 90 0318 3.

3. Composition of foods, Australia. Volume 3 Dairy products, eggs and fish (1990). AGPS Canberra. Cat. No. 90 8034 9.

4. Composition of foods, Australia. Volume 4 Fats and oils, processed meat, fish and vegetables (1990). AGPS Canberra. Cat. No. 90 1154 4.

5. Composition of foods, Australia. Volume 5 Nuts and legumes, beverages and miscellaneous foods (1990). AGPS Canberra. Cat. No. 90 2430 1.

6. NUTTAB90. Nutrient data table for use in Australia (1990). AGPS Canberra. Cat. No. 90 1429 8 (nutrient data on diskette).

7. English, R., Lewis, J. (1991). Nutritional values of Australian foods (1991). AGPS Canberra. Cat. No. 91 0212 3.

8. FOODfiles; Datafiles of the New Zealand Food Composition Database, Version 2.0 (1991). FOODDATA, Department of Scientific and Industrial Research, Palmerston North.

9. FOODsearch; Applications software plus datafiles of the New Zealand Food Composition Database, Version 2.0 (1991). FOODDATA, Department of Scientific and Industrial Research, Palmerston North.

10. NZ Diet/1; Applications software plus datafiles of the New Zealand Food Composition Database, Version 1.0 (1991). FOODDATA, Department of Scientific and Industrial Research, Palmerston North.

11. Milligan, G.C., Webster, D.W., Burlingame, B.A. (1988). The New Zealand Food Composition Tables, FOODDATA Biotechnology Division, DSIR Palmerston North

12. Visser, F.R., Burrows, J.K. (1983) Composition of New Zealand Foods 1. Characteristic Fruits and Vegetables, DSIR Bulletin 235.

13. Visser, F.R., Hannah, D., Bailey, R.W. (1990) Composition of New Zealand Foods 2. Export Fruits and Vegetables, Crown Copyright, DSIR Biotechnology and Royal Society of New Zealand.

14. Visser, F.R., Gray, I., Williams, M. (1991) The Composition of New Zealand Foods 3. Dairy Products, Crown Copyright, DSIR Crop Research and Dairy Advisory Bureau.

15. The New Zealand Food Facts (1990). Department of Health and DSIR, Mills Publishing.

16. Vlieg, P. (1988) The Proximate Composition of New Zealand Marine Finfish and Shellfish, Crown Copyright, DSIR Biotechnology Division.

17. Klensin, J.C., Feskanich, D., Lin, V., Truswell, S., Southgate, D.A.T. (1989). Identification of Food Components for INFOODS Data Interchange. United Nations University, Tokyo.

18. Klensin, J. (in press). INFOODS Food Composition Data Interchange Handbook. United Nations University, Tokyo.


Australian publications and datafiles are available for purchase from the Australian Government Publishing Service Mail Order Sales, GPO Box 84, Canberra ACT 2601, Australia.

New Zealand publications and datafiles are available for purchase from FOODDATA, Department of Scientific and Industrial Research, Private Bag, Palmerston North, New Zealand.

Various South Pacific publications are available freely or for purchase from the South Pacific Commission, B.P. D5, Noumea Cedex, New Caledonia.

# Using European Data - An Interview with Dr. Lenore Arab-Kohlmeier

**John C. Klensin**
**INFOODS**
**Cambridge, Massachusetts**

Using European Data
An Interview with Dr. Lenore Arab-Kohlmeier

**Background:** Dr. Arab-Kohlmeier was originally expected to attend the conference and participate in the "Using International Data" discussion. When an unexpected schedule conflict arose, she was interviewed over the telephone by Dr. Klensin, the session moderator. That interview was recorded and played at the conference. Because Dr. Arab-Kohlmeier's comments contain an important perspective and information, the interview was transcribed at appears below. The text was only very lightly edited after transcription to preserve the flavor of a very informal and unprepared telephone conversation. Consequently, the remarks have not been as intensively reviewed by either participant as ones in a formal paper would have been.

**JcK:** What we are trying to do is increase some general understanding of the use of international data, that is data that originated in one country for local use and then is to be used in some other country. To some degree the intra-European data issues reflect most of the problems although you benefit from some commonalities of food habits. In addition, beginning recently, you have had the advantages of pressures originating outside strictly scientific community to consolidate and compare those data; i.e., with the EC situation. The nature and limitations of the NDB audience is such that discussions and examples that relate to import of data into the U.S.A. will be most understood and meaningfully related to, with issues of use of U.S. data outside the U.S. being next most comprehensible. To some extent this is an examination of what resources are available and where one should look for information. So, I am particularly interested in finding out how to locate and, when necessary, assess the usefulness of, available data in terms of your directory and its updates, the INFOODS directory and critical national or regional databases that don't appear in those directories. You mentioned to me in Crete, for example, that for any sort of official uses in Germany, the Souci-Fachmann-Kraut book is basically not used, and that the BLS data base is used instead. To what extent is it generally true through Europe, that there are things that people think are the tables and things which people use that are different?

**L A-K:** O.K., I'll try to address that. But, first, John, you say hello to all of the people that I would have missed seeing at the meeting and tell them that for any more details, of course, I am accessible as is Clive West and other individuals who are involved with the Eurofoods operation in Europe--in Wageningen and in Berlin.

The BLS (Bundeslebensmittelschlussel) is a nutrient data base with values on 12,000 foods and recipes that people can get from us on tape. It is a combination of many different food composition tables and a tremendous amount of guesstimate. The user has a database with no missing values and all of our users are coming up with comparable results at least as far as the nutrient values are concerned in the epidemiological studies conducted in Germany and Austria. This experiment began five years ago, in the meantime, the Ministry of Research and Technology and the Ministry of Health in Germany support only studies that use this common database in an allowed version. The users sign a contract saying they will not change the values in the data base. This is for the very simple purpose of comparability. The BLS is used in German-speaking countries. There are 200 groups using it now, and it is based largely on Souci-Fachmann-Kraut but also on other tables.

That tells you about the German-speaking area. We're talking about 100 million people in Central Europe. Now, in other countries in Europe, there are national and local systems being used for different purposes. Probably in Holland and in Denmark, the systems are most centralized. There had been attempts to have a national nutrient database that is distributed widely and which is getting feedback from the users in terms of what they need and what they are missing. These are generally not data systems which are estimating missing values. And, although we have come a long way in the last ten years in knowing what different users in different countries need, the availability of data bases to users from other countries is not yet a reality. I cannot tell you how many databases are actually up and being used.

We have more or less an overview of the printed books, and those printed books are not that which is commonly being used. It's rather the other way around: people look for the software first, the software is the leading edge in data access. Users tend, if they can find software in their own language, to look for inexpensive software that is easily usable, widely available and accessible on PC's. If they find that in another language, they'll borrow and then translate it, rather than try to set up a system new. They take less into consideration the concept that "maybe those nutrients don't apply to my food". Now, in Europe, and especially as we're racing towards 1992, that becomes much less of an issue because we are sharing food, and we will be sharing food more and more.

**JcK:** A consolidated table like that with a lot of missing values in it is presumably subject to all of the advantages of a consolidated table (whichever one is using) and all the disadvantages of something which mixes analytic data from a number of sources and non-analytic data from a number of sources. Has that been a problem, or is the consistency for your application really more important than that set of issues?

**L A-K:** Really, consistency is more important. We're talking about the use in the epidemiological context where our first area of concern is the measurement error associated with portion sizes. We are investing much more in trying to reduce that than in trying to have more harmonious tables. So, to the earlier question, I guess the basic question is, "are food tables becoming generic?" and I would think the answer is "yes": it is not a situation in which more and more resources are being poured into getting more analytical data. It's rather the other way around. We are heading more and more towards having a European database, and as a matter of fact, there are some different groups within the European Community--in EUROSTAT, in DG12, in the FLAIR group, and in the Medical Research Division--that are becoming increasingly interested in having some common European food nutrient tables for cross-country analyses. I don't know if these attempts will be successful, but we should know within the next two to three years.

**JcK:** How does one go about getting these kinds of tables and tapes which are essentially undocumented in terms of the published references?

**L A-K:** That's a good question.

**L A-K:** Our initial attempt to try to answer that question was in putting together that book called *European Food Composition Tables in Translation.* For each country we identified in 1987 the table or the various tables that were being published nationally, and who was producing them, and how many versions they had produced. In addition, we tried to get, or create, translations of the introductory materials into English, so that foreign users could also, in a reasonable way, begin using this "foreign" data, which may or may not have been "foreign" once you started looking at what was actually coming up between the lines. And certainly, that resource is still appropriate because the people producing tables have not changed, but it is outdated in that there are new tables that have come up since then. And, it seems the data--food nutrient composition tables in book form--are being produced more and more rapidly. At least in the Scandinavian countries, they are coming out every year, or every few years, with new national tables. So, it's hard to get an overview of what is there unless you are directly corresponding with the institutes or the working groups responsible for producing them.

**JcK:** But that publication, like the INFOODS directory for other areas, was much, much more successful, I think, in identifying printed materials than it was identifying the data bases. As we are seeing a shift toward data bases and software instead of the printed tables for most serious uses, that becomes more and more of a problem.

**L A-K:** There is no way to overview these currently. There is, as far as I know, no way to get at the successful software databases. They are being produced partly by some industries. UNILEVER, for example, has one or various that they are using and Milupa has some as well, but there is no European analogy to your Nutrient Databank Conference where people are coming together saying, "here's what we have and here's what we are using." What we're attempting to do in Berlin in the context of our WHO Collaborating Center on Nutritional Epidemiology is to install electronic mailing lists where people can ask those kinds of questions and those people doing nutritional epidemiology in Europe can address them. Of course, we have no constraints on Americans joining these mailing lists. I know you are on our mailing list (NUTEPI, see at end for instructions), so you could transfer this information further: how to get on this group, and what the node and access name is.

**JcK:** If one is trying to combine tables or pull data from some tables into other tables, what do you see as the special and hard problems these days? For example, in building a European table and pulling French data into a German table, or any of those data into a U.S. table?

**L A-K:** There are such a range of problems, and they haven't changed very much in the last decade. The first problem is really one of translation and definition of foods. I have personally given up on the idea that we are going to identify identical foods--the best we can do is identify similar foods--between the tables. In most cases, one will never have complete documentation on where those foods actually came from, or where the analytical data came from. Though it might be subconsciously troubling to many people, we have seemed to learn to live with it. If there is a problem, it is an illness in all the food compositions tables. There was an attempt in 1983 and 1984 to see if merging tables was technically possible. That was very successful, but was then laid aside with the final report on that project to the European community. Currently, in IARC (the International Agency on Research in Cancer), this task is being re-begun with more recent food nutrient composition tables from seven different countries, in the context of a large cohort study on diet and cancer. A cohort of 400,000 people in seven European countries is now being put together, and the need for a common database is very evident and has a top priority. We will be meeting and looking at this nutrient database which the group of Elio Riboli in Lyon is producing the week after next, actually concurrent with your nutrient database meeting.

**JcK:** How much U.S. data is being borrowed for use in Europe. Do you have any sense?

**L A-K:** There was a project that attempted to look at that very specifically, and I believe that the rough estimates were between 40 and 60 percent of the total mass of data that was available in Europe was actually borrowed from USDA. Now, if that had been stable at that point, or if it is decreasing, I really couldn't tell you, but there is a tremendous amount of American data in the European system, I know that the Greeks, for example, are relying very heavily on the USDA data, and within our own Souci-Fachmann-Kraut tables, there are 20 to 40 percent borrowing from the basic nutrient data, from various versions of USDA. This does not always mean that is the most recent USDA data either. It is from different versions of different handbooks. What one needs, or what one would like is a Michelin guide for composition tables, and we don't have anyone putting that together right now.

**JcK:** What one should both have the Green guide which tells you how to navigate and the Red guide which tells you something about the quality of it.

**L A-K:** Exactly.

**JcK:** Do we have any sense as to how much European data is trickling into databases in the U.S.?

**L A-K:** I would not have any idea.

**JcK:** I would assume some of McCance and very little of anything else, but you don't have any sense either?

**L A-K:** I wouldn't have any overview of that from this end of the Atlantic.

**JcK:** O.K. well sometimes you have insights, so I thought it was worth asking. I think that about takes care of the questions. Thank you, and on behalf of the Conference, our thanks and regards, and wishes that you could be joining us.

**L A-K:** Well, thank you as well, and I wish the Conference success. I have enjoyed them very much when I've been there, and since I was at the last one in San Francisco, I particularly regret not being able to be there this time. We will, in Berlin in October of next year, have a meeting of our Collaborating Center on Nutritional Epidemiology addressing the issues of validation and quality control which might be an occasion for participants of the meeting to cross the other way and ask the questions directly, and we are inviting you all with open arms to see the new Berlin at that point in time. This will be October 1992.

**JcK:** I was about to ask how people get information on that conference, but it sounds like we can announce it at next year's Nutrient Databank Conference.

**L A-K:** Or, they can write to me to be put on our WHO Collaborating Center Newsletter where all the information on European meetings is contained.

Address:

Dr. Lenore Kohlmeier, Bundesgesundheitsamt, Postfach 330013, 1000 Berlin 33, Germany

Information for subscription to "NUTEPI" electronic mailing list:

Send mail to LISTSERV@db0tui11 (from BITNET) or LISTSERV@tubvm.cs.tu-berlin.de (from Internet, UUCP, etc). The first, and only, line of the message should read:

SUB NUTEPI Your Name

Where "Your Name" is your actual first and last name.

**References:**

Arab L, Wittler M, Schettler G. European food composition tables in translation. Berlin: Springer-Verlag, 1987.

Heintze DM, Klensin JC, Rand WM, International Directory of Food Composition Tables, 2nd ed, Cambridge, MA: INFOODS, 1988.

West EC, ed. Inventory of European Food Composition Tables and Nutrient Database Systems. Uppsala: National Food Administration, 1989.

# Some Issues and Problems in the Usefulness of Chemical Composition Data Across Boundries

**Ricardo Bressani**
**LATINFOODS Coordinator**
**Guatemala, Central America**

Due to the increasing trend in food imports and exports, in food aid programs, in preparing and consuming foreign food dishes, in studies on the relationship between diet and disease, in nutrition intervention projects, in nutritional assessment of specific populations, in food labeling and food fortification and other nutrition-related activities, food composition data from various countries should be more useful to local and foreign needs. However, the use of foreign chemical composition data presents many problems due to the complexity in developing and compiling data for food composition tables. The problem becomes more difficult when the data available is relatively old as is the case for information from most of Latin America. No efforts have been made to upgrade the quality of the data in spite of the fact that today there are more analytical results on many foods that were available in the past when the tables were first compiled.

In this short presentation it will be indicated what are some of the issues and problems which may be encountered in the use of food composition data from Latin America and maybe attempt to provide solutions. However, before discussing that, I like to indicate that in technologically advanced countries, diversity in nutrient composition values for a single food is very small, while in the less technologically advanced countries diversity in composition values of the same food is quite wide. This is, therefore, a very important issue which must be kept in mind.

## AN OVERVIEW ON PRESENT FOOD COMPOSITION DATA IN LATIN AMERICA

Food analysis in most countries of Latin America was initiated some 40-50 years ago for the purpose to study the nutrient intake of rural and urban populations and to have additional criteria for the evaluation of the nutritional status of the population at large or of specific population groups. Although many kinds of foods were analyzed, the number of samples per food were small, because of the urgency for the use of the data at the time. The nutrients analyzed included proximate composition, calcium, phosphorus and iron, carotene and Vitamin A, $B_1$, $B_2$, Niacin and Ascorbic acid. Not many efforts to upgrade the quality of the data and to increase the numberof nutrients per food have been made since 1960, when most Tables were published and some activity was initiated since 1983 when the concept of LATINFOODS was introduced (1). Present available tables are in general, relatively good but most are far from being the kind of document which is needed now and more so in the future, at the local level and for use in other countries. Therefore, there are limitations in the analytical data which make the process of data interchange difficult with a need to proceed with caution.

## SOME OF THE PROBLEMS

The main issues and problems a person who would like to use food chemical composition data will find can be included into anyone of four groups (Table 1): Group 1 is related to the problems which arise due to insufficient sample identification, sample classification and sample description. In Group 2 of issues and problems one may include aspects related to the food sample itself. It is well known that chemical composition of vegetable and animal products is highly

**Table 1 - Issues which may cause problems in the interchange of chemical composition of foods**

| | |
|---|---|
| Group 1 | Sample identification, classification and description |
| Group 2 | Issues specific to the factors influencing the chemical composition of the food |
| Group 3 | Processing factors - Industrial, Home |
| Group 4 | Sample Analysis - Units of Expression |

related to genetic factors, environment and cultural practices, postharvest technology activity and marketing. Processing techniques and conditions represent a 3rd group of factors. The same kind of food may be prepared at home and at the industrial level, and steps in the process may be different in the two situations. Group 4 of factors include all activities associated to the actual chemical analysis, which begins with the selection of a representative sample, an appropriate number of samples, sample preparation and preservation, method of analysis, standards used and expression of the analytical value, both in term of specific moisture basis and the units used (2,3,4).

## SAMPLE IDENTIFICATION, CLASSIFICATION AND DESCRIPTION

The problems associated with the identification, the classification and the description of the foods apply to all food groups, but probably more so for those of vegetable origin. Possibly this is one of the most important problems which causes difficulties even within closely related regions and more so between different regions. Before describing the problems it is important to define the terms identification, classification and description (Table 2).

Identification is used as a means to include the food into its specific food group. It should provide also a common name and synonyms. The name in other languages is also very useful. Classification is the term used to provide the scientific name and its taxonomic identity. Description is used to indicate the part of the food used whether from vegetable or animal origin, the method of processing, the type of meat cut, whether the food is consumed fresh or cooked, the ingredient composition and amounts in mixed dishes, to mention a few. To further clarify the value and significance of these terms the following examples are given: common beans are identified as food grain legume and classified as *Phaseolus*, while maize is a cereal grain belonging to *Zea*. Although there may be problems in identification and classification, in general these issues will not give much of a problem, which if they exist they can be relatively easy to solve. Sample description is very important, but it is more complex and only minimal information is often available. For example there are light and heavy creams, sweet and sour creams, but only cream is given in the Latin American table. The problem here is that food standards and definitions are not available and if they are, the standard is not implemented. Processed cheeses and other types of cheese may have well-known names such as cheddar, pecorino, mozzarella, but due to the variability in milk composition and to the process, composition of the product is not alike to the true kind of cheese or standard of other countries. Standards of identity are usually established from chemical composition data in which the chemical composition of the raw material plays a very important role, as well as the process, which is useful in adjusting the food to the standard. Often the standards of identity established in any country for a processed food are taken from the standards of identity of another country without taking into consideration basic differences in chemical composition of the raw material.

**Table 2 - Definition of Terms**

| | |
|---|---|
| Identification | Type of Food; Food Group; Common Name and Synonyms |
| Classification | Scientific Name; Taxonomic Identity |
| Description | Part of food Consumed; Percent utilization; specific process used to make it. Single food - Dish |

The importance of the significance of these terms is greater when mixed dishes rather than single foods are examined, which required more thought and evaluation before the food values are to be used. For example, tacos, enchiladas are different in Guatemala from those from Mexico from those in the U.S. It is important to provide ingredient composition. Ideally the tables orother document should contain a Food Description Dictionary, which would be of significant help in the application of food composition values, by persons using data from other countries and who are not familiar with the food. An example for three foods are given in Tables 3, 4 and 5.

These food descriptions can be very useful and may be drafted in different ways to provide the user with as much information as possible. One possibly would be a description based on the food chain, so that all descriptions of foods follow a standard system of description.

**Table 3**

*Tortillas de maiz* - (Maize tortillas). A food purchased in the market or made at home from industrial flour or from hard and soft whole maize often white by cooking maize at boiling T in a 0.6-1.0% lime solution based on maize weight for 50-65 min, which removes the seed coat, followed by washing, grinding, dough prepartation and baking of 20-50 g protions made into flat cakes on a clay surface sprayed with thin lime eater solution and heated up to 210°C for up to 2.5-3 minutes on each side, to be conusmed warm, reheated or dried.

**Table 4**

*Frijoles fritos* - (Fried beans). Black or red beans are cooked in water until soft. They are then ground and strained with the cooking liquor. The strained beans without most of the seed coat are cooked with 20-25% oil addition and onions to yeild a thick soup-like product, which with futher water evaporation yeild a pastw with a shape. They are produced industriallly but seed coats are not removed. The product's preparation differs between countries.

**Table 5**

*Atol or Atole* - A generic name for a thick drink, served hot, commonly made from maize at the dough-stage by water addition and pressing to produce a milk-like liquid which is then cooked with milk and sugar to a thick consistency. Often anatto flour is sprayed on the top and 3-5 cooked black beans are added for appearance. It can also be made from other gelatinized cereals, lime-treated maize dough roasted food legumes and cereal/oilseed flour mixtures.

These food descriptions are useful in the analysis, selection and acceptance of analytical values. In the case of maize tortillas one would expect a high calcium level since the process use lime, which may also carry with it other minerals. One may expect low levels of B-vitamins because cooking is done at alkaline pH, and dietary fiber may be lower than in raw maize since the seed coat is removed.

The fried beans industrially produced may have higher levels of dietary fiber than the home-made product because bean seed hulls are not removed. It should be indicated also that fried beans as prepared in other countries may be prepared in a different way.

Finally, with atole, protein and calories, as well as other nutrients will be different if made with milk rather than with water, of if it is made from roasted legume grains or flour blends of maize/oilseeds.

Descriptions are applicable only to foods which receive a type of process not common in foreign countries. Foods which are commonly found and consumed in a number of countries need only their true identification and classification if ingredient composition and process are alike, and in minimal description.

The identification, classification and description of foods will facilitate data interchanging and the process of developing a coding system equal or similar to others already available and also the means to use it in computer programs (4,5).

## ISSUES AND PROBLEMS RELATED TO THE SAMPLE

As it was indicated before, present chemical data in the Tables was obtained some yearsago. Since then important changes have taken place in all aspects of the food chain leading to the production, availability and transformation of the food to be consumed. For example even though a number of improved varieties of maize and of beans have been introduced in agricultural production, there are many land races still being produced. The same applies to other food crops. Likewise, agricultural practices have improved but traditional practices still persist. The same applies to animal food products. In rural areas for example chicken meat and eggs came from unidentified breeds which feed themselves with what they are able to find in contrast to chicken meat and eggs marketed in large cities which are produced in large scale with compounded feeds. Chicken meat in rural areas comes from chickens more than twice as old as chickens grown on balanced feeds. About 85% of the swine population in Latin America is represented by native swine much different in composition than swine of improved breeds. Feeding practices are also different leading to differences in chemical composition. Most fruits and vegetables are transported without much protection and remain under natural conditions for a few days until consumed. They may be harvested and marketed at different stages of physiological maturity. There are seasonal effects, milk during the dry season (no rain) is more concentrated from cows fed dry and high fiber grass, while milk from the rainy season is more diluted, due probably to the high water content of the fresh grass being consumed. With respect to cattle, meat cuts are different and not only are breeds different, but all meat comes from animals which have been grass fed from 2.5 up to 4 or 5 years of age.

On top of this, one must add the activities of post harvest technology which have changed throughout the years. Therefore, present day food intake studies for estimation of nutrient intake using chemical data that is relatively old runs the risk to give untrue estimations. These problems, however, can be solved if the samples were to be described as indicated before which will help in selecting the chemical values to be used. A further problem is that many of the chemical analysis are on raw samples, and only a few values are available on mixed dishes.

## ISSUES AND PROBLEMS RELATED TO PROCESSING

One of the main problems in the Latin American Food Tables is that about 85% of the results are on raw natural foods, and it is well known that processing affects chemical composition with some processes more than others. Processing may not change much the content of macro nutrients, but significant changes take place in the micro nutrients. The problem is much more complex because the same food processed at home may have a different composition as compared to the food prepared at the industrial level, and even at home level, in different regions from the same country, the same food may be processed differently or with different ingredients resulting in different values (Table 6). This diversity is large enough in single foods, and significantly larger in mixed foods, an example of which is the "Tamal". This is made from maize, rice or potatoes, may have different types of meat cuts, with or without dried fruits, with and without chili sauce, tomatoes or different levels of fat, from vegetable oil to animal fat.

## ISSUES AND PROBLEMS RELATED TO THE CHEMICAL ANALYSIS

There are a number of issues in this group of factors which are of interest. These are the selection of a representative sample, the number of samples analyzed, the preparation of the sample and its preservation, the analytical techniques and standards used and the expression of the value. The representativeness of the sample may be a problem, since the same material to be analyzed may come from a great many number of environments and subjected to many factors which may or may not affect the content of the nutrient to be analyzed. In a country like Guatemala, for example, there are many types of maize being grown, and it is a dangerous situation to use the analytical values obtained on only one or two types. The same applies to other foodstuffs. Many foods are grown at different altitudes over sea level in different seasons. Not all nutrients are affected by these environmental factors to the same extent. Macro nutrients possibly less than micro nutrients. Another problem is the number of samples analyzed which for most foods is usually only a few, probably not more than 5. Exception are some basic foods (Table 7). Again, there are more samples analyzed for macro nutrients as compared to the number samples for the micro nutrients. Sample preparation is also a problem, particularly when the sample is not well homogenized beforeportions are weighed for analysis. With respect to the analytical technique used, this is probably the least of the problems, although there are examples where the analytical method gives conflicting results such as carotenes. However, it may be a factor. Finally the chemical value is expressed on the basis of its actual moisture content, and in many cases the moisture value is not provided.

**Table 6 - Differences In Ca, Fe and Zn Content of Maize Tortilla by groups of Households In two regions In Guatemala**

| | North Region** | | South Region*** | |
|---|---|---|---|---|
| Mineral* | Average | Range | Average | Range |
| Ca | 202±74 | 99±476 | 217±41.5 | 167-250 |
| Fe | 2.7±0.8 | 1.9±6.4 | 7.0±4.8 | 4.0-16.0 |
| Zn | 3.4±0.1 | 2.1±4.4 | 5.4±0.4 | 4.8-5.7 |

* mg/100g d.w. basis
** Krause, 1988
*** Bressani et al., 1988

**Table 7 - Number of Analysis on Single Foods (Latin America Table)**

| Food | Macro Nutrients (No. Of Analysis) | Micro Nutrients (No. Of Analysis) |
|---|---|---|
| White Rice | 32-36 | 20-33 |
| Maize | 51-53 | 50-53 |
| Arepa (from maize)-threes countries | 1-4 | 1-4 |
| Tortilla (from maize) | 3-25 | 3-25 |
| Chad | 65-70 | 17-70 |
| Amaranth | 7 | 4-7 |
| Spinach | 1 | 1 |
| Common beans (P.vulgaris) | 133-270 | 165-270 |
| Bananas | 1-3 | 2 |
| Guanabana (Sour sop) | 8 | 5-6 |
| Guava (Psidium guajava) | 25-27 | 21-37 |

Reference (6).

## A PROCESS FOR THE SELECTION AND USE OF CHEMICAL DATA

It is difficult to propose a standard way or guidelines to select chemical data from a table of one region to be used in a situation in another geographical region. However, there are some steps which can be followed, besides those already discussed. An example may help. Potassium analysis, for example, were requested for vegetables, fruits and root crops from one geographical region to be used for nutrition studies in another geographical region. The steps to be followed include:

1. It is important to identify, classify and describe the vegetables, fruits and root crops of interest in both geographical regions to learn if they are the same.
2. A decision has to be reached if the genetic characteristics, the agricultural practices followed in their production, and handling practices are the same in both geographical regions.
3. Knowledge must be available on how the foods are consumed (fresh or processed). The method of processing used for consumption in both regions must be established.
4. It is important to determine if the chemical analysis was done on the fresh or processed sample. If more than one processes, or a different process is used, the analysis must be on each or on the one used in the region where the values are to be used.
5. A reference to the method of analysis must be given and an explanation should be provided on the units used for expression of the results.
6. Once the above sequence of analysis is finished the values provided could be used.

## THE IMPLICATIONS OF THE USE OF DATA WITH RESPECT TO APPLICATION PROGRAMS

It is obvious that for all applications of food composition data, these should be the bestwhich can be obtained, however, there are various applications which can be met relatively well with what it is available in terms of data quality. For example in studies of nutrient intake based on food intake data, analytical food values may be not the best, since food intake data is much less accurate than most analytical values, even with dietary surveys where foods are weighed. On the other hand, specific nutrient intervention programs would require the best values available, as close to the situation as possible. The best values would be needed to know what is the present intake of the nutrient in question, as well as to know what levels to test in the intervention. Other applications such as food fortification or metabolic studies would require the best available values. Likewise, studies between diet and disease would require also the best possible values for a better analysis of cause and effect relationship. Many nutrition intervention studies have failed to support metabolic nutrition results because knowledge has been lacking on the actual nutrient intake of the baseline situation and because the analytical values of the foods at the baseline situation were not selected adequately.

In conclusion, the available analytical values of the foods of one region can be used in another region if there is a process of selection of the values taking into consideration for the selection all possible information on the sample. The use of foreign data can be enhanced if there is increased communication between countries with more advanced expertise in the problem with those with less experience.

## REFERENCES


1.LATINFOODS. Memorias de la Primera Reunión sobre Tablas de Composición de Alimentos LATINFOODS. 11-14 November, 1986. Arch. Latinoamer. Nutr. 37(4):609-810, 1987.

2.BRUCE, A. AND L. BERGSTROM. User requirements for Data Bases and applications in nutrition research. Food & Nutr. Bull. 5:24-9, 1983.

3.BRESSANI, R. The data required for a Food Data System. Food & Nutr. Bull. 5:69-76, 1983.

4.PALACCHI, W. Standardized food terminology: an essential element for preparing and using food consumption data on an international basis. Food & Nutr. Bull. 8:66-68, 1986.

5.KLENSIN, J. C. Interchange issues in Data handling. pp. 42-49. In: Proceedings of the First ASIAFOODS Conference. Eds. W. Rand, A. Stuckey, A. Valyasevi & K. Tonsisirin. The Institute of Nutrition, Mahidul University, Bangkok, Thailand. 17-21 September, 1984.

6.INCAP-ICNND. Tabla de Composición de Alimentos para Uso en América Latina. INCAP, Guatemala. June 1961.

7.KRAUSE, V. M. Rural-urban variations in lime maize consumption and the mineral content of tortilla in Guatemala. Center for studies of sensory impairment, aging and metabolism. Guatemala City, Guatemala & School of Dietetics and Human Nutrition, McGill University, Montreal, Canada, 1988.

8.BRESSANI, R., M. BREUNER & M. A. ORTIZ. Acid- and Neutro-detergent fiber and mineral elements content of maize and tortilla. Arch. Latinoamer. Nutr. 39:382-391, 1988.

# Issues and Problems in Using Food Composition Data in Asia

**Prapasri Puwastein, Ph.D.**
**Mahidol University**
**Nakhon Pathom, Thailand**

## INTRODUCTION

Food is a major component of man's environment. Data on actual nutrient composition of foods are critical for the important activities of a great variety of individuals and groups including those involved with epidemiological research into disease patterns, formulation of dietary recommendations, health assessment of individuals and populations, and national and international trade in foods. Owing to the fact that food types and their preparation vary tremendously from country to country and region to region, efforts have been made by each country in the generation and compilation of food composition data specific for their own use. These data, moreover, can be beneficial nationally and internationally when required.

However, preparation of good quality data with complete information which are accessible and meet users' needs requires many resources including time, planning, management skill, manpower and finances. Limitations and constraints in this process introduce a number of difficulties for users which must be overcome before any true thought can be given to data and information transfer. It is hope that continuous sharing of problems and experiences among generators, compilers and users as being done in this conference can lead to solutions to some of the most perplexing problems and provide a working concept rather than an abstract principle.

This presentation highlights a small sample of persisting problems and as reported by ASIAFOODS members which continue to influence the quality of data. All of these areas determine our ability to realistically exchange and use food composition data cross-nationally through common system.

## NATURE AND PROBLEMS OF ASIAN FOOD COMPOSITION TABLES

For the Asian region, existing FCTs can be divided into roughly two types (**Table 1**). The first type includes those in which all data are generated from local resource materials, while the second includes local data as well as those from neighboring countries, other regions or even international data when available. According to relatively recent directories and country reports (1-3), most countries in the Asian region have their own national FCTs, though a regional FCT has not as yet been formulated although a definite desire exists among ASIAFOODS members.

**TABLE 1: CHARACTERISTICS OF AVAILABLE FCTs IN ASIA**

| Country | Type of FCTs A | Type of FCTs B | Language used Native | Language used English |
|---|---|---|---|---|
| Brunei | ------No national FCT available------ | | | |
| China | o | | o | (1) |
| Indonesia | o | ? | o | |
| Japan | o | | o | (2) |
| Korea | | o | o | |
| Malaysia | o | | | o |
| Myanmar | | o | ? | ? |
| Philippines | o | | | o |
| Sri Lanka | | o | | o |
| Taiwan | | o | o | |
| Thailand | o | o | o | o |

(A) : Generated by their own resources.
(B) : Generated by their own resource and compiled from other FCTs
? : No information is available

(1) : FCT is being translated.
(2) : English data is partially provided.

Unfortunately however, these FCTs are not without their problems, and these are thought to be mainly due to the generated FCTs themselves. Overall, the most common obstacles are the timeliness of data (are they up-to-date?), completeness (what information and/or data are missing or inadequate?), users' intentions, and the extent to which foreign data is included in local/national FCTs without regard to varying cross-national considerations. In particular, food composition data are vital to the work of a wide variety of users. This diversity, in itself, breeds problems, because data users have different expectations and requirements, many of which are beyond the capabilities of the generators. In addition, generators oftentimes possess very limited resources. Mechanisms to solve this problem are desperately needed, and which will most likely entail bringing users, generators and compilers into closer contact, so they can work together in a more coordinated fashion. Other specific problems are as follows.

**1. Completeness of Available Food Composition Data in Asia: Specific Cases**

*1.1 Nutrient Components*

The completeness of the data in Asian FCTs varies from country to country. As shown in **Table 2**, most include at least fifteen basic nutrients and some cover many more. The general information available are: 1) main nutrients (proximate composition) in which dietary fiber was presented as crude fiber; and 2) certain minerals and vitamins, i.e., Ca, P, Fe, vitamin A, carotene, vitamin B1, B2, niacin and vitamin C. Some particular nutrients can be obtained from certain FCTs or research papers which were published in local or international journals (i.e., macro-minerals - Na, K, Mg; trace elements - Zn, Cu, Mn, Se, I; and some vitamins - B6, folic acid, B12; as well as fatty acids and amino acids).

At present dietary fibre data, which is becoming more and more essential for clinical and epidemiological studies, are missing in all FCTs in Asia. However, limited data on dietary fibre content in foods -- as reported by Japan, Singapore and Thailand -- can be obtained from published journals or scientific papers (3-6). Developing dietary fibre data (total and its fractions, to be included into the national FCTs) is in the future plan of many countries, i.e., Japan, Malaysia, the Philippines and Thailand. An INFOODS recommended method for determination of dietary fiber and its fractions is requested.

For fat soluble vitamins, more accurate estimation of vitamin A value in a number of local fruits and vegetables, as well as foods of animal origin, are being carried out in Malaysia and Thailand using high pressure liquid chromatography. Hopefully, updated data on this nutrient will soon be available for inclusion in FCTs. Apart from vitamin A and carotene, only Japanese FCTs contain data on vitamin D and vitamin E. No data on vitamin K is available in all FCTs.

The Philippines FCT not only contains additional data on Mg, Cu, Zn, I, Mn vitamin B6, B12 and folate, but they also provide some data on anti-nutritional factors, e.g., phytin and oxalic acid as related to the bioavailability of nutrients.

*1.2 Food Items*

Data on nutrient composition of many food items are missing in each national FCT, especially foods infrequently consumed, specific local foods such as unconventional (but seasonally common) protein sources of people in rural areas, wild foods, and newly manufactured and/or distributed foods, either raw or processed. A crucial consideration which must be made in data exchange is to what extent are certain types of data wanted on an international scale. For example, the Institute of Nutrition has nutritional data on edible insects used as foods in rural Thailand. But whether or not this data is desired by INFOODS or other network members, and how it is to be compiled for dissemination, is unknown.

Many FCTs present data of raw foods even for foods usually consumed after being processed. Different processes of food preparation prior to consumption (such as boiling, steaming, frying, roasting) affect various nutrients at different degrees. This effectively limits the use of cross-national data which does not indicate exact preparation methods. Due to differences in culture and food availability, the process of cooking foods and ingredients used vary from country to country or even at different regions in the same country, which in turn greatly affects the composition of the processed foods. Many data on nutrient composition of processed and traditionally cooked foods can be obtained from FCTs; cases in point are the FCTs of Malaysia and the Philippines which also notes with serving size, and that of Malaysia also includes ingredients and methods of preparation/cooking. However, use of data on cooked foods can mostly be used within respective countries due to specificity of recipes and cooking methods.

**2. Data Presentation**

*2.1 Presentation of Analytical Data*

How data are presented will determine utility amongst local and international users as well as amenability for computerization. To be of maximum convenience, worldwide generators including those in Asia, express their analytical data in terms of weight of nutrient and by weight of sample, for instance, ug/100g, mg/100g, g/100g edible portion depending on the concentration of each nutrient in the respective food sample. Among the 4 FCTs available in English at the ASIAFOODS regional Center, the Philippine FCT provides data of processed foods not only per 100 g edible portion but also per household unit. This type of data expression as well as those for fresh and cooked foods are in need of specialized users such as community nutritionists and epidemiologist for dietary assessment. The Malaysian FCT also gives information about serving size (in weight) of traditional cooked foods which could facilitate local users in assessing dietary intake.

**TABLE 2: AVAILABLE NUTRIENT INFORMATION IN DIFFERENT FCTs IN ASIA**

| | China | Indonesia | Japan | Korea | Malaysia | Myanmar | Philippines | Singapore | Thailand |
|---|---|---|---|---|---|---|---|---|---|
| **Main nutrient** | o | o | o | o | o | o | o | o | o |
| DF | | | o(1) | | (2) | | (3) | o(1) | o(1) |
| **Minerals** | | | | | | | | | |
| Ca, P | o | o | o | o | o | o | o | o | o |
| Na | | | o | o | o | | o | o | o(1) |
| K | | | o | ? | o | | o | o | o(1) |
| Mg | | | (2) | ? | | | o | | o(1) |
| Fe | o | o | o | o | o | o | o | o | o |
| Cu | | | (2) | | (2) | | o | | o(1) |
| Zn | | | (2) | | (2) | | o | | o(1) |
| Se | ? | | (2) | | o | | | | o(1) |
| I | | | (2) | | | | o | | (2) |
| Mn | | | (2) | | (2) | | o | | |
| **Vitamins** | | | | | | | | | |
| Vit A | o | o | o | o | o | o | o | o | o |
| Carotene | o | | o | ? | o | | | | o |
| RE | | | o | | o | | | | o(1) |
| Vit D | | | o | | | | | | |
| Vit E | | | o | | (2) | | | | |
| Vit K | | | | | | | | | |
| Vit B1 | o | o | o | o | o | | o | o | o |
| Vit B2 | o | | o | o | o | | o | o | o |
| Niacin | o | | o | o | o | | o | o | o |
| Vit C | o | o | o | o | o | | o | o | o |
| Vit B6 | | | (2) | o | | | o | | |
| Vit B12 | | | (2) | o | | | o | | o(1) |
| Folate | | | (2) | o | (2) | o | | | |
| Panto. | | | | o | | | | | |
| FAs | o | | o | o | (2) | | (3) | (2) | o(1) |
| AAs | o | o | o | o | (2) | | o | | o |
| **Others** | | | | | | | | | |
| Choles. | o | | o | o | (2) | | o | (2) | o(1) |
| Phytate | | | | | | o | o | o | o(1) |
| Oxalate | | | | | | | o | | |
| NaCl | | | | o | o | | | | |
| Total | 19 | 14 | 26 | >20 | 21 | 12 | 29 | 19 | 29 |

o : Nutrients included in FCT
? : Uncertain as to inclusion in FCT
(1) : Data available in published or unpublished papers, not yet included into the National FCTs
(2) : Data are being developed
(3) : Future plan for analysis

*2.2 Language Used and Food Identification*

Native languages are used in some of the national FCTs in Asia such as those of China, Japan, Korea, and Indonesia. International information exchange in these cases will be impossible unless they are translated into English. Presently, China is translating its FCT with assistance from the Institute of Cancer Research, the National Institute of Health, USA. On the other hand, Japan can only partially provide English data. Thailand's FCTs present part of "information to the users" in her native language, but nutrient composition data are presented in English. Malaysia and the Philippines present information to the users, names of foods together with the local names and nutrient data in English. Malaysia's FCTs are even more convenient since the English name is first presented, followed by the Bahasa Malaysia name.

Translation of local food names into English sometimes causes errors. In different parts of a country, colloquial food names may be different and can be misleading, which causes difficulties when translation is made. In addition, some foods may have different names in different countries or regions, or different foods may even have the same name. Standardization of food names is a very tedious work, however efforts should be made, in cooperation with taxonomists of plants, animals, insects, etc., to correctly identify the analyzed foods. Photographs of specific local foods or food models would also assist in this process, though they have limited computer application. To facilitate data exchange for common food items, the nomenclature should be standardized by INFOODS in cooperation with counterpart members.

### 3. Availability of Related Information

Whenever an FCT is generated, it is essential to document carefully the factors involved which can and do contribute to the variability of the data. Users can then be made aware of potential problems surrounding the data, and they can judge for themselves as to whether or not the data are of sufficient quality for their use, or how they can work around the problems.

Regarding the availability of communication information in Asian FCTs, those of Malaysia and the Philippines are good representatives. The information given in their FCTs include:

1) Explanation notes on the FCTs
   - grouping of foods
   - nomenclature and description of foods
   - numbering of food items
   - source of data
   - methods of sampling and number of sample to be analyzed
   - methods of nutrient analyses
   - converting factor for protein calculation from nitrogen content
   - specific factor used to calculate energy
   - specific physiological energy factors for calculating the caloric value of foods (the Philippines's FCTs)
   - Recommended Dietary Allowances
   - abbreviation
2) Description of cooked foods which include
   - name of foods
   - serving size
   - ingredients
   - methods of preparation /cooking
3) Index of English names of foods
4) Index of native or local names of foods
5) Index of scientific names of foods
6) Bibliography

### 4. Computer Facilities and Systems in Asia

Ultimately, users need more than just data; they require the "hardware" and "software" to interact with these data. In several Asian countries (e.g., Japan, Malaysia, the Philippines, Thailand), computerization has assisted in decreased the time users require in painstaking, often routine, calculations, data generation, revision, compilation and management. This has also led to the more effective use of available data by a variety of users. The computer software are either self-developed or modifications of commercial programs such as FCODA (food composition data base system developed by the National Food Research Institute at Tsukuba, Japan), NUTRITIONIST III, FRAMEWORK, dBase III through IV, Foxbase, Lotus 1-2-3, etc. But the most important need is for a software "package" which is not complicated, flexible and allows the users to expand and revise nutritional information. It should also facilitate users in maximizing their ability to interact with available data.

## RESPONSES, SUGGESTIONS AND REQUESTS FROM FCT USERS REGARDING DATA USAGE

Asian countries which generate their own FCTs via their own resources always use their own data. "Imported data" are used only for cross-checking their data when doubts about reliability or validity arise, or as source of nutrient information when the required data is missing in their own FCTs. Nonetheless, a number of problems are reported in using and comparing imported data in Asia. Some of the most important include the following, and, once again, these are not new problems. The presentation of users' responses to the problems are also included, however alternatives responses are expected from the group discussions at this conference.

### 1. Missing Data and Data Inconsistency

While data on some specific nutrients or data on newly manufactured foods and foods infrequently consumed are often missing from national food tables, there is a major and crucial gap regarding the composition of foods as they are consumed. Specific to this problem is the issue of "raw" versus "cooked" foods.

*1.1 Raw Foods*

Using data from imported FCTs whenever nutrient or food items are missing in the national FCTs is a common answer from Asian users. The three most frequently used FCTs **(Table 3)** are the USDA Agricultural Handbook, McCance and Widdowson's *The Composition of Foods* as revised by Paul and Southgate and FCTs for use in East Asia, FAO (1972). One mai reason for using these FCTs is due to the English language presentation provided as well as their global distribution. Actually, using FCTs from neighboring countries is the most practical and desired, if the needed data are available and in English. Among FCTs available in Asia (**Table 2**), those of Japan, the Philippines and Thailand contain a greater amount of nutrient data (26, 29 and 29 nutrient items, respectively). Additional data regarding such aspects as dietary fiber, trace elements, and cholesterol are to be included in these national FCTs. In the mean time published data are available on request.

*1.2 Cooked Foods*

Missing national FCT information about cooked foods is another major problem encountered by users. As mentioned earlier, the nutritive value of cooked foods can vary greatly due to differences in the ingredients and raw materials used, size of raw material, methods of cooking, time and temperature used, etc. The best way to handle this problem is to gather representative samples of the cooked foods and assess them for the desired component. However, this option requires resources that users rarely have available. Therefore, most of the users compute the nutritive values of cooked foods by themselves as based on the nutritive value of raw ingredients that available in the FCTs. In order to gain more representative data by this option, careful step by step trials should be followed and appropriate considerations made in reporting the steps and standardizing data for computerization.

**TABLE 3. INTERNATIONAL FOOD COMPOSITION TABLES USED AMONG ASIAN COUNTRIES**

| COUNTRIES | USDA | EUROPE | FAO | OTHERS |
|---|---|---|---|---|
| Brunei | | o | | Malaysia |
| China | o | o | o | |
| Indonesia | ----------No information available----------- | | | |
| Japan | o | o | | |
| Korea | o | | o | NIH, Japan |
| Malaysia | | | o | |
| Myanmar | | o | o | India |
| Philippines | ---------No information available----------- | | | |
| Sri Lanka | | o | o | India |
| Taiwan | o | | o | Japan |
| Thailand | o | o | o | Philippines, Japan, Malaysia Australia |

USDA : USDA Agricultural Handbook
EUROPE : McCance and Widdowson's *The Composition of Foods*, revised by A.A. Paul and D.A.T. Southgate, 1978.
FAO : Food Composition Table for Use in East Asia, FAO, 1972.

*Computation of "Raw" versus "Cooked" Foods*

There are two ways to compute the nutritive value of raw versus cooked foods.

**Method 1.** *Calculation from Raw Food Recipes*

*Step 1.* Development of specific conversion factor to reconvert weight of cooked ingredients to raw ingredients. Different methods or conditions of cooking the same food contribute different conversion factors. This step is quite tedious, time and budget consuming. The factors also cannot account for the heat-labile nutrients, i.e., some vitamins. However, the conversion factors obtained are very valuable and can be used effectively to improve the quality of data in dietary assessment.

*Step 2.* Carefully weigh cooked ingredients of a representative dish to be estimated.

*Step 3.* Convert weight of cooked ingredients to raw ingredients using the specific conversion factors of specific ingredients.

*Step 4.* Calculate nutritive value of the cooked dish from available nutrient composition data of raw materials. The estimation of cooked food is then computed

**Method 2.** *Calculation from Computed Cooked Foods*

*Step 1.* Standardization of the ingredients by cooking the dish of interest several times. Each time the ingredients are recorded and average weight of raw ingredients as well as the obtained cooked food are estimated. If standard recipes of the interested dish are available, for example, from a dietary survey, by weighing method or from a standard cookbook, the weight of ingredients can be directly used.

*Step 2.* The estimation of nutrient composition of the cooked foods in a particular portion is then computed from the available data of raw materials in FCTs.

*Step 3.* The nutritive value of the same food (cooked) with different weights as consumed by a subject can be then estimated from the computed nutritive value of cooked food obtained in step 2.

Both methods are based on the assumption that there is a loss of only water during different processes of cooking or processing. However, in some foods especially fatty foods, not only water is affected by cooking, but lipid content and other nutrient may gain or loss. These methods cannot correct for the nutrients gained or lost by the process.

1.3 *Effects of Seasonal Variation and Planting Areas on the Nutritive Value of Foods as Arising from Inconsistent Data*

A question of whether seasonal variation as well as differences in planting areas affect nutritive value of foods is a major concern of users. These two areas arose since food composition data in international FCTs are often not consistent when compared to one another. At least in the first case, and possibly the second, this will affect the construction of FCTs and their ultimate cross-cultural utilization either through printed or computerized documents. Up to now, no data on these aspects is available in Asia.

**2. Lack of Data Expression per Portion Size**

In the process of a typical dietary assessment, a 24-hour recall method is used wherein the amount of food consumed by a subject is recorded as portion size using standard spoons or cups. Since most of the food composition data is expressed in terms of weight of nutrient per 100 g of edible portion, information on portion size is rarely given. Consequently, many users have difficulties in estimating the nutritive values from local and imported FCTs. Measuring the size of foods as well as weighing total and edible portions of a food sample prior to analysis by the generators is requested from the users. Nevertheless, using of a standard set of food models with known edible portions should be the right direction to this problem. The food models do not only help the subjects to recall how much food they have consumed, but they also help the interviewer in estimating the exact amounts of edible portions consumed. This in turn increases the quality of dietary assessment data.

## DISCUSSION AND CONCLUSION

The problems encountered in the FCTs of differing Asian countries, as well as those associated with data imported from other sources, are not new problems; they were discussed in detail at an earlier 1987 conference report (7) as well as communications between ASIAFOODS network members. Yet each is a problem which must individually be corrected if good quality data is to emerge.

### Mutual Awareness

People in Asia (including ASEAN) and many developing nations, who are working with the diverse aspects of food composition data, are not strongly aware of their efforts' similarities and the issues of which they must address. This has led to a persisting tradition of independent, non-coordinated activities resulting in the duplication of efforts, data incompleteness, and incompatibilities in data exchange. For many foods, only a limited number of nutrients have been analyzed, and data is still missing for many food items. In addition, many crucial data sets are scattered among institutions without any serious effort being made to compile them.

In not every case is this a problem. Where resources have permitted, data generation, the construction of adequate FCTs (e.g., Philippines, Malaysia), and to a small extent data exchange have progressed. Yet what is persistently missing is a sense of community within the field which would avoid problems of resource wastage and increase the likelihood of compiling more reliable international directories of food composition. Strongly recommended is the development of a mechanism for increasing awareness and cooperation among institutions and member countries within the field.

### Limited Resources

To improve food composition data generation and compilation, all involved institutes require intensive efforts and resource use (e.g., time, manpower, finances). Very often, limited numbers of well-trained staff may be available for the required work programs. Yet since the latter are not "research-oriented", funding is difficult to obtain. These constraints can easily limit activities in up-dating and up-grading national FCTs in Asia. However, strong coordinated efforts could assist, to some extent, in overcoming these constraints.

### Limitation of Standard Guidelines

Increased awareness exists regarding the need for a systematic program for generating, compiling and disseminating food composition data. Many standards and guidelines remain desperately needed in a number of areas if data exchange is to become a reality. Among ASEAN countries, for example, a draft INFOODS Manual, prepared by Greenfield and Southgate in 1987 (8) was distributed. This systematic guideline contains information of the production, management and use of food composition data, and it is often used by many ASEAN countries. Unfortunately, not all Asian countries have access to this guideline, largely because it is not a formal publication as yet. Nonetheless, use of such a guideline to effectively improve the quality, quantity and accessibility of food composition data is strongly recommended.

### INFOODS Active Working Group

It goes without saying that the crucial leader in solving persistent problems in Asia and elsewhere is INFOODS itself. But the latter need not bear the entire burden. The construction of a **INFOODS *Active* Working Group**, comprised of key resource persons from each sub-network (e.g., EUROFOODS, OCEANIAFOODS, ASIAFOODS, NORFOODS, LATINFOODS), could assume this responsibility and work towards seeking support to correct some of the needs mentioned above, and which to one extent or another persist within each sub-network. Suggested activities for the Working Group are as follows:

1. The development of a standard questionnaire and the implementation of a survey on certain nutrient and food item data should be among countries within the region.
2. The formulation of a standard guideline for generating specific data, of special importance is cooked food.
3. Due to inconsistency in generated data, the working group could coordinate interlaboratory trials using an agreed upon procedure or guideline.
4. Another crucial area is the standardization of food nomenclature.
5. Provide guidelines as per the basic, necessary information to be included in national FCTs.

## REFERENCES


1. Proceedings of the first ASIAFOODS Conference, Bangkok, THAILAND: Country Report Summary , 1984:52-84.

2. International Directory of Food Composition Tables, INFOODS, First edition, January 1986.

3. Proceedings ASEAN Workshop on Food Data System, Bangkok, THAILAND: Country Report, 1989:98-162.

4. Candlish JK, Lynn G, Lee HP. Dietary fiber and starch in some Southeast Asian fruits. J Fd Composition and Analysis 1987;1:81-84.

5. Candlish JK, Gourley L and Lee HP. Dietary fiber and starch contents of some Southeast Asian vegetables. J Agric Fd Chem 1987;319-321:1987.

6. Puwastien P, Valaipatchara U, Kongkachuichai R. Dietary fiber in common Thai foods. J Nutri Thai Assoc 1991 (in press).

7. Rand WM, Windham CT, Wyse BW, Young VR. Food Composition Data : A User's Perpective. Food and Nutrition Bulletin Supplement 12, UNU, 1987.

8. Greenfield H, Southgate DAT. Guidelines to the Production, Management and Use of Food Composition Data System. An INFOODS Manual (draft), 1987.

# Using International Nutrient Data

**Lena Bergstrom**
**National Food Administration**
**Uppsala, Sweden**

## USING INTERNATIONAL NUTRIENT DATA

Foods, products and dishes in the Nordic countries - Denmark, Finland, Iceland, Norway and Sweden - are very similar, if not basically the same. But of course there exist some differences especially with respect to pure national food items such as Swedish fermented herring, Icelandic geo-thermal water and Finnish Karelian pastry.

As a result, you might believe that our data of food nutrients would also be similar, but that is not always the case.

The food composition data may differ depending on e.g. the following fac- tors: natural variations, standards, enrichment or fortification, analyses, energy factors, dietary fibre and recipe calculation.

### Natural variations

Danish studies show considerable differences in nutrient contents between herring from the North Sea and the Baltic Sea, depending on the fish food available. Similar differences in nutrient contents could probably also appear in other wildlife animal species.

### Standards

Some food items in the Nordic countries are standardized. Let us just takeone example: low-fat milk. You must be aware that the fat content may vary. In Sweden low-fat milk contains 0.5% fat, in Finland 1% and in the other Nordic countries 1.5% fat.

### Enrichment or fortification

Enrichment of food items should also be taken into consideration. The low-fat milk can serve as an example again. This milk is enriched with vitamins A and D in Sweden and Finland but not in the other Nordic countries.

Another example is common wheat flour. In Iceland and Sweden this flour is available both plain and enriched with thiamin, riboflavin, niacin and iron, and in Sweden also with vitamin B6. In the other three countries the wheat flour is unenriched.

The enrichment of a food item is usually indicated in the food description, but, as in the Swedish base, enrichment nutrients are not indicated.

But there are also other types of enrichment to be considered, that is soil and feed enrichment. In Finland e.g., due to the very low content of selenium in the soil, the addition of selenium to fertilizers was begun in 1984. These fortified fertilizers will affect both plant and animal foods, and the outcome will be an increase of the selenium intake in humans.

### Analyses

In the Nordic countries the co-ordination and evaluation of analysis methods have been arranged to provide for reliable common methods. This work is carried out by the Nordic Committee on Food Analysis. As a result of this co-operation, the differences in analysis results depend more on natural food variations and sampling than on methods.

One mode of analysis procedure, sometimes used in the Nordic countries, may create discrepancies. That is when, e.g., one year you analyse the proximate constituents for a specific group of foods, and the following year you analyse the fatty acids, vitamins or minerals for the same food group, but most probably do not use the same samples as for the proximates.

### Energy calculation

In Sweden we calculate energy in kJ according to the European Economic Communitys directive on nutrition labelling for foodstuffs.

The energy value to be declared shall be calculated using the following conversion factors:

| | | |
|---|---|---|
| cabohydrate (except polyols) | 4 kcal/g | 17 kJ/g |
| polyols | 2.4 kcal/g | 10 kJ/g |
| protein | 4 kcal/g | 17 kJ/g |
| fat | 9 kcal/g | 37 kJ/g |
| alcohol | 7 kcal/g | 29 kJ/g |
| organic acid | 3 kcal/g | 13 kJ/g |

As the other Nordic countries use the conversion factor 38 kJ/g for fat, there are some differences in the energy values for food items containing fat. Moreover, three countries use 30 kJ/g for alcohol.

Some Euoropean countries do not follow these EEC factors, e.g. Great Britain and France who use 3.75 kcal for carbohydrates. One Swedish firm exported some products to England and the British firm changed the Swedish energy values. The firm contacted us and asked what to do. I informed the firm that in England they use 3.75, and I suggested that the producer write to the importer and ask if that was the reason for changing the energy values. A prompt answer came: "Yes, we use 3.75 for carbohydrates, probably because we are British".

## Dietary fibre

Dietary fibre holds a unique position with regard to energy. Suggestions on conversion factors have a range of 2 - 5.4 kcal/g. A mean of 2 or 3 might be adequate. But according to the EEC regulations on foodstuffs, dietary fibre should not be calculated for energy. The food manufacturers in the Nordic countries follow the EEC regulations, and the compilers of food composition tables and bases will also most certainly do that. A solution to the problem of dietary fibre and energy could be, e.g. in food composition tables, to give two energy values for foods containing dietary fibre, one without and one with energy for fibre. This procedure is now being used in the Finnish food composition tables.

## Recipes

The people responsible for the nutrient data banks in the Nordic countries have somewhat different approaches to calculation for nutrients in recipes. Yield factors in cooking are commonly used, but the factors may differ for the same dishes. Factors for nutrient changes, mostly for vitamins, are used differently. In Norway, vitamin B factors are used in calculating bread recipes. In Denmark, factors for different food groups and cooking methods are created, and in Finland, the emphasis is put on the cooking methods in constructing factors. We, in Sweden, have up to now been using loss factors for 5 vitamins in all recipes of our standard recipe file. In the future we hope to coordinate the recipe calculation further.

## Compiling data

Within Norfoods (the Nordic group working on compatibility of Nordic food composition tables and nutrient data banks), we have agreed that, when nutrient values are missing, we should primarily use Nordic analysis data. But how often have we not been grateful to the Americans, who have analysed so many food items for so many nutrients. Of course we are also much obliged to other countries for their analysis work on foodstuffs. Without international assistance and cooperation, we compilers would be hard pressed to produce complete national food composition tables and nutrient data banks.


**References**

Ahola M, Bergstrom L, M'ller A, Reykdal O, Rimestad A. H: Activities of Norfoods 1987-1988. 4th Eurofoods Meeting, May 31-June 3, 1989. Proceedings, 13-19. National Food Administration. Uppsala, 1990.

Bilde B, Leth T: The Danish Food Monitoring System. Status after the first 5-year period. 4th Eurofoods Meeting, May 31-June 3, 1989. Proceedings, 109-129. National Food Administration. Uppsala, 1990.

Council Directive of 24 September 1990 on nutrition labelling for foodstuffs. (90/496/EEC). Official Journal of the European Communities, No L 276/40, 6.10.90.

Goranzon H: Energiinnehll i odigererbara kolhydrater. Nringsforskning 35: 48-49, 1991.

Jokela K, Bruce: Selenberikning av handelsgodning i Finland. Vr Foda 37: 287-289, 1985.

M'ller A (ed.): Levnedsmiddeltabeller 1989. Nutrient Composition of Danish Foods. Stork'kkencentret, Levnedsmiddelstyrelsen. S'borg, 1989.

Nordic Committee on Food Analysis. Booklet 1991. Available from Ulla Edberg, National Food Administration. Uppsala.

NORFOODS Report 1988:1. Nutrient calculation systems in the Nordic countries. National Food Administration. Uppsala, 1988.

Rastas M, Seppnen R, Knuts L-R, Karvetti R-L, Varo P (eds.): Ruoka-aineiden ravintoainesislt÷. Nutrient composition of foods. Publications of the Social Insurance Institution. Helsinki, 1989.

Reykdal O (ed.): Islenskar Naeringarefnat÷flur. Icelandic food composition tables. Volumes 1 and 2. Reykjavik, 1988.

Statens Ernaeringsrd: Matvaretabell. (Norwegian food composition tables). Landsforeningen for kosthold & helse. Oslo, 1984.

Statens Livsmedelsverk: Livsmedelstabeller. Energi och nringsmnen. (Swedish food composition tables). Statens livsmedelsverk. Uppsala, 1986,revised 1988.

# Using International Data: Conclusions

John C. Klensin[1]
INFOODS Secretariat
Cambridge, Massachusetts

Data from across national and regional boundaries have been important for some time, and are becoming more so. People travel and migrate and bring their food styles with them to countries where those foods are neither well-represented in tables and databases nor easy to reproduce exactly with locally-available ingredients. However, considerable effort is required for the intelligent use of data in one country that were developed in another. There are difficulties in the identification of foods[2], in cooking techniques and methods used to estimate or impute values for mixed foods[3], in conventions about nutrients[4], and in conventions and methods of reporting and presenting information[5].

The presentations in this session examined issues in data originating outside the United States in various regions of the world. They identified, although to different degrees, several common obstacles to data availability and use. These included:

**It is still very difficult to locate data.**

Finding data of interest depends to a great degree on personal contacts. Detailed directories published in various parts of the world and the overview produced by INFOODS cover tables and other printed materials much better than they cover actual databases. The comparative listing of software systems and associated data formerly produced annually for this conference has no equal elsewhere. In general, one must find an expert in a particular country or region and ask her or him what is available or for additional people to contact.

**There are consistency problems with names, food identification, and nutrients.**

No "system" or "magic bullet" will make these difficulties disappear, since they result in large measure from real differences in foods and scientific theories and assumptions. However, for some types of problems and uses, the importance of precise identification and even precise values may have been significantly exaggerated. More research is needed on the topic of "how much difference does it make?".

These are particularly important considerations for systems intended to precisely name, and thereby identify, foods. At the best, these systems provide only partial solutions to very complex problems. For other problems and approaches, they may be largely irrelevant. The field has not yet established a critical literature of evaluations, especially comparative evaluations, of naming and description systems.

**Choices of software have come to dominate choices of databases.**

Where data are processed with computers, people tend to choose software that is attractive from a human interface and language standpoint as long as required data seem to be available. Less attention is paid to issues of data quality. Similarly, for at least some purposes, consistency of the results of calculations is seen as more important than the accuracy or representativeness of those calculations. These approaches can pose particular problems when data are compared across national or regional boundaries, since they often result in a reduced availability of the kinds of descriptive materials needed to evaluate inter-database or inter-table consistency. The field still lacks a literature of critical evaluations of databases and, especially, software packages and a consensus about how to perform such evaluations.

**International coordination is getting better in some places, worse in others.**

INFOODS is increasingly operating on the regional basis which was intended. Several of the regions are very strong and have active programs. INFOODS works with them to facilitate inter-regional activities and coordination and to provide technical advice and, as a result, has fewer programs of its own. The lack of a functioning regional organization for North America may eventually pose problems for people in this region trying to work with data from outside of it.

While the availability and use of "international" data can be very helpful, and is sometimes vital to an adequate understanding of nutrition or health status, obtaining and using such data in a scientifically-responsible way is difficult. INFOODS and its associated regional groups continue to be very active in this area. There are gradual improvements as a result of these, and other, efforts that actually focus on international issues, rather than expecting all other countries to conform to one's own local preferences and needs.

[1]INFOODS Secretariat, Room N52-457, Massachusetts Institute of Technology, 77 Massachusetts Avenue, Cambridge, MA 02139.

[2]Truswell AS, et al. "INFOODS Guidelines for Describing Foods: A Systematic Approach to Describing Foods to Facilitate International Exchange of Food Composition Data". *J. Food Composition and Analysis* 4, 1 (March 1991), pp. 18-38 provides an international perspective on this issue.

[3]See Rand WM, et al. *Compiling Data for Food Composition Data Bases.* In press, United Nations University. To appear late 1991.

[4]Klensin JC, et al. *Identification of Food Components for INFOODS Data Interchange.* Tokyo: United Nations University, 1989.

[5]Klensin JC. *INFOODS Food Composition Data Interchange Handbook.* In press, United Nations University. To appear late 1991.

# Database and Programming Needs of Research/Health Care Dietitians

**Phyllis J. Stumbo**
**University of Iowa**
**Iowa City, Iowa**

This morning I want to depart from our discussion of databases and consider the front-end program that manipulates the data. I am particularly interested in the programming needs of health care or clinical dietitians in their role of nutrition counselor. This morning I will narrow my focus to the printout that results from nutrient calculations.

The printout is one outcome of computerized nutrient calculation programs often used by nutrition counselors. While a nutrition counselor knows that attitude, not knowledge, is the primary force that drives eating habits, counselors should be sure their clients have adequate knowledge to support making desired changes. The computer can help provide accurate information about dietary intake and is a potentially useful information resource in the counseling process.

The two reasons clinicians might use nutrient intake data in counseling are to identify problems, and to monitor dietary change. Both objectives involve collecting information, sorting and interpreting the information, and giving it back to the client in a more meaningful form.

Information about diet is most meaningful when it represents a whole day or an average day's dietary intake. A variety of software is available to help develop and interpret this type of dietary information. The interpretation may involve food groups or exchanges or a detailed report of nutritive and nonnutritive components. Some programs manipulate over 50 components, and since a typical intake includes at least 10 food items, the computer will generate 500 or more data points for each daily intake. This is more information than clients can easily remember. The client seldom wonders how much zinc or boron is present in each food but rather will ask "How good is my intake?" or, "What is wrong with my diet?"

To evaluate how well computer programs provide information for nutrition counselors and help to answer client's questions I surveyed software developers to obtain a printout for evaluation. I polled producers of the 52 diet analysis programs housed in the National Library of Agriculture in Beltsville, MD asking for printouts that evaluated the nutritionally good and nutritionally poor menus I provided. Table 1 lists the seventeen developers who responded with printouts or a copy of their program (33% of those surveyed).

After receiving program output I was faced with the problem of how to evaluate the results. Some of the best ideas for evaluating printed media come from the newspaper and magazine publishing field where attention is given to appearance and readability and from education where attention is given to setting objectives and guiding the learning process.

Programs should give attention to layout and design, focus attention on important information and avoid a cluttered appearance. Printouts should present selected information, but not too much information. The amount of information can be controlled by setting an educational objective and selecting information to develop that objective.

My request to program developers surveyed was for printouts to support nutrition counseling in general rather than to identify a specific objective. Today's discussion is limited to one set of printouts, whereas the actual software may provide a variety of other options. I will limit my discussion to printouts describing the "good" menu which consisted of:

Breakfast:
- 1 cup oatmeal
- 1/2 cup 2% milk
- 1 slice wheat bread
- 1 Tbsp jelly
- 1/2 cup frozen orange juice

Lunch:
- 2 slices whole wheat bread
- 1/4 cup tuna salad
- 1 medium pear
- 1/2 cup low fat cottage cheese
- 12 oz can cola

Dinner:
- 1 chicken breast, broiled
- 1 medium baked potato
- 2/3 cup steamed broccoli
- 2 tsp tub margarine
- 1 cup iceberg lettuce
- 1/2 medium tomato, fresh
- 1 Tbsp French dressing
- 1 cup skim milk

Snack:
- 1 fresh apple
- 8 vanilla wafers

**Table 1 - Developers who provided program output**

1. CBORD Diet Analyzer, The CBORD Group., 61 Brown Road, Ithaca, NY 14850$995.
2. DAP, Dietary Analysis Program, USDA-HNIS, Belcrest Road, Hyattsville, MD 20782 (sold by NTIS) $60
3. DAS, Dietary Assessment System, Softech Computing Co., 2401 Hirschman Lane, Hartland, WE 53029 $195
4. Dietician, Alsoft, Inc., POBox 927, Spring, TX 77383-0927 MAC $59
5. DINE, DINE Systems, Inc., 586 N. French Road, Ste 2, Amherst, NY 14228 $170 - $345. IBM, MAC
6. Eat for Health, Genesee Intermediate School District, 2413 West Maple Avenue, Flint, MI 48507-3493 $25
7. NDS, University of Minnesota, Nutrition Coordinating Center (NCC), 2221 University Avenue, SE, Ste. 310, Minneapolis, MN 55414 $7,000
8. Nutri-Calc Plus, Camde Corporation, 4435 S. Rural Road, Ste. 331, Tempe, AZ 85282 $225
9. Nutri-Tally, Nutrition Counseling, 221 Seventh St., N. Columbus, MS 39701 $125
10. NAS-2, Nutrient Analysis System 2, DDA Software, PO Box 26, Hamburg, NJ 07419 IBM, MAC $289.95
11. Nutriplanner, Practorcare, 10951 Sorrento Valley Road, San Diego, CA 92121 $800
12. N3 and Rite Byte, Nutritionist III, and Right Byte, N-Squared Computing, 3040 Commercial St., SE, Ste.240, Salem, OR 97302 N3-$495, Right Byte- $595
13. PRUCAL, Dept of Food Science and Human Nutrition, Colorado State University, Fort Collins, OR 80523 $98
14. Professional Dietitian, Wellsource, PO Box 569, 15431 SE 82nd Drive, Ste.D, Clackamas, OR 97015 $495
15. Sante, Hopkins Technology, 421 Hazel Lane, Ste. 300, Hopkins, MN 55343-7117
16. The Good Health and Diet Program, Diet Research, Inc., 3665 Brighton Way, Reno, NV 89509 $60 Negotiable
17. You Are What You Eat, Marshware, PO Box 8082, Shawnee Mission, KS 66208 $41.95

The educational objective of the majority of printouts I received was to depict overall nutritional adequacy. The programs gave information to describe the nutritional value of the menu in a variety of formats. If I had any criticism it is they tried to give too much information on one or two pages. There seem to be two main messages in all the programs, guided by both the Recommended Dietary Allowances (R.D.A.) (1) and Dietary Guidelines (2). The first is avoid deficiency (fulfill the R.D.A.) and then second is avoid excess (follow the Dietary Guidelines). The mistake too often made is to try to make this a single message.

Graphs communicate information in a memorable way. Most programs used graphs to convey information about dietary adequacy and dietary excess. I will show you a few examples to illustrate how computer programs express nutritional adequacy and excess. By far the most attractive result was from Nutrient Analysis System 2 (NAS-2) by DDA Software shown in Figure 1. I'd Like to point out a couple of factors that I think make this a good display. First, the evaluation shows nutrients up to 100% of the R.D.A. and not beyond. Nutrition experts do not recommend exceeding the R.D.A. so depicting more than 100% does not enhance the message on adequacy. If the message is avoid deficiency, then this graph explains clearly how well this diet avoids deficiency. Each bar would extend to the right hand margin if the diet provided 100% of the RDA for all nutrients.

CBORD and NCC take the RDA graph to 200% (Figure 2) On casual observation, this graph seems to say there is too little of some nutrients and too much of others, i.e., it is not 'balanced'. That is the wrong message for this diet. While vitamins C and E exceed the RDA there is no danger to this level of intake but graphing it to 200% dwarfs the other nutrients. As a nutritionist I am not concerned about excesses until they reach several times the RDA. If I wanted to wa… a client about excess nutrient intake I would probab… want to show intakes of 5 to 25 times the RDA as … sometime consumed in *supplements*, but *pref…* food graph to show the value of meals m… 3 to 100% of the R.D.A.

Returning to the NAS-2 graph, (Figu… that what I thought was a well selected m… than 1/2 the R.D.A. for zinc. Follo… nutritional dogma I chose tuna fish fo… chicken for dinner -- no red meat. Therefo… contained only 1/2 the RDA for zinc. It is …

for dietaries to be below the R.D.A. for zinc, in fact it is quite common. I lecture to dental students and review their computer assignment of evaluating a colleague's diet using the USDA's Dietary Analysis Program. They are almost always zinc deficient by their computer analysis, and it is not because the data base is sparse.

We know the RDA's tend to err on the high side. The 1990 RDA's reduced the folacin allowance by half, and suddenly we are no longer a nation deficient in folacin. Although some of my colleagues disagree, I believe the RDA is too high for zinc and will eventually be reduced when we have sufficient data to make a reduction. If I used dietary analysis routinely in practice for the purpose of evaluating nutrient adequacy I personally would want to remove the zinc comparison for my clients when I think showing this so-called deficiency would be a disservice. Therefore I would want to create customized bar graphs for my use.

For example, in an intervention study involving children we use diet calculations to track cholesterol intake. Figure 3 shows a graph we routinely create on "Cricket Graph" to illustrate this one point. I would like to have more options in nutrient analysis programs to create customized printouts which use graphs such as this. A realistic educational objective for a client may involve only one piece of information. Comprehensive bar charts are too complex for some counseling sessions.

The third dietary guideline reads "Choose a diet low in fat, saturated fat, and cholesterol" which is defined as 30% or less of calories from fat and less that 10% from saturated fat. It is difficult to translate this into a useable guideline for clients. It is like trying to hit a moving target. If you have a meal with 45% of calories from fat and you add 3 slices of bread, the percent (since it is part of the whole) could drop to 20% for a small meal or to 40% for a large meal.

Programers typically illustrate this guideline with a pie chart. Two years ago Chor San Khoo, speaking about food labels, said "No one understands pie charts", and I tend to agree (3). At least I am not skillful at making sense out of pie charts as they relate to dietary macronutrients. The best use I have seen for pie charts illustrating nutritional concepts was on the back of a Healthy Choice Frozen meal. Figure 4 illustrates three values from a frozen meal label, 2% of calories from saturated fat, 17% of the maximum 300 mg cholesterol, and 11% of the maximum 2400 mg sodium. Notice that these pie charts are limited to one value. Pie charts depicting percent of carbohydrate, protein, and fat and sometimes alcohol try to give too much information. I find it hard to attend to differences in four factors in these 2 pie charts from the NAS-2 program shown in Figure 5. Usually the important piece of information that can be clearly depicted is the percent of calories from fat and since that is the only percent expressed in the seven dietary guidelines, a customized bar chart showing percent of calories from fat would be meaningful and memorable.

Figure 6 is a graph from the Rite Byte program which illustrates another problem with nutrient calculation programs. Note that zinc is deficient, and so is saturated fat. Mixing RDA and dietary goals introduces a mixed message. I doubt that any dietitian would suggest a client should increase saturated fat to reach the recommended 10% of calories or to increase fat intake that is greater than 20% up to the recommended 30% of calories, but putting them on the same chart suggests to me that 100% is the goal and ideally all components should be increased to meet the 100% goal. In fairness to Rite Byte notice that "+" before some components indicate they are dietary goals and not RDA's, but combining goals and R.D.A. in this way requires attention to each separate detail rather than providing the opportunity to make a global assessment based on how complete the graph appears.

Some programs go beyond the RDA and dietary guidelines. DINE has taken on the difficult task of interpreting data that is not always best when it is 100% by using heavy lines to indicate the recommended range for a whole host of factors. The DINE graph in Figure 7 shows the macronutrients, and the division between animal and vegetable protein. The first bar is for protein showing 19% of calories is from protein.

**Table 2. Polyunsaturated fatty acid (PUFA) content of 7 food items from two nutrient calculation programs.**

| | PUFA, g | |
|---|---|---|
| | NDS | DAS |
| Oatmeal | 0.87 | - |
| WW Bread | 0.32 | - |
| Jelly | 0.01 | - |
| WW bread | 0.65 | - |
| Tuna salad | 3.65 | - |
| Cola | 0.07 | - |
| Vanilla wafers | 0.77 | - |
| Total | 6.33 | 0 |

The recommended percentage is 10-15% shown by the dark line. Total fat is 22% shown with up to 30% recommended as illustrated by the dark bar extending to the 30% mark. Saturated, monounsaturated and polyunsaturated fatty acids are each less than the recommended 10% as illustrated by heavy lines. The diet also falls short of the recommended 50% of carbohydrate as complex carbohydrate, and plant protein falls far short of the recommended 50%. This is a complex message presented graphically. There are many ways to depict data and the graphs displayed here each tell the nutrition story in a unique way.

Nutri-tally has a different way to graph data uing a circle with 3 sections where the inner circle represents 100% of RDA. Figure 8 from Nutri-tally pictures all the information in one graph. While it depicts a lot of information in one place it has the problem of accentuating excesses and the format is not as familiar as either the bar or pie charts, so for me it requires more study to fully understand the message.

Before leaving the topic of printouts I feel obligated to at least mention the data. I said I would not talk about databases, but you really cannot separate the report from the data. Missing values is a common problem among several programs. For example the PUFA in the DAS printout was 10 g rather than as much as 16 g in calculations from other programs. DAS provides no PUFA values for the 7 food items listed in Table 2 whereas NDS's more complete data base listed values for each. Total PUFA would have been at least 1 1/2 times greater in the DAS calculation if all foods actually containing PUFA had complete fatty acid information.

Even with a complete database, providing accurate nutrient information is a challenge for the clinical dietitian. I provided to software developers what I thought was a reasonably detailed day's menu composed of common foods about which data is available, but the resulting nutritional data still showed inconsistencies.

Some inconsistencies occur when conflicting decisions are made by coders. For example, coders differed on what they assumed a serving of chicken breast was. I provided only as much information as I thought would be available from a client, and since my clients who eat chicken usually describe white meat as a chicken breast, even though 1/2 breast is the most common cut in the market and cafeteria. Most coders did assume a chicken breast serving is 1/2 breast, but two coders assumed a chicken breast serving was the whole breast. This type of discrepancy could be attributed to untrained coders or poor documentation. Figure 10 depicts the discrepancies in caloric value for the sample menu made by several programs with the two higher values resulting from choosing a whole rather than a half chicken breast.

This points out the importance of training coders and standardizing procedures and terminology. Inadequate training is a problem as is the proliferation of sparse data bases. Last year I issued a plea for programs to suppress summations of nutrient columns containing vacant data cells - where no data is available for a nutrient thought to be present in a food, to try to prevent use of incomplete data (4). As a result of this year's project I received a new program from Hopkins Technology called Sante which thrilled me by offering a choice of suppressing totals from columns with missing values. It was refreshing to see a program actually announce that the database was sparse, and to offer a way to deal with the sparseness. Here, I thought, was a reasonable solution, I could still have incomplete totals, but since I have to ask for them I should at least be aware of their weakness. I was mildly disappointed that this held true only for totals displays, and summations of incomplete values were still displayed when a line-by-line data table was generated. At least with the line-by-line printout I can see that some values are missing, but too many users do not consider this deficiency and even when programs alert users to incomplete data many users are not discouraged from reporting the results.

Summary: There is a rich resource of commercial computational programs that manipulate nutrient data available for purchase. The primary educational objective inherent in those I reviewed is overall nutritional adequacy. Assessment of macronutrient percentages is commonly used to evaluate fat intake and comparison of nutrient totals to the RDA are the most prominent educational goals.

We need better ways to individualize programs for unique counseling needs. This is possible to some degree with most of the programs reviewed. N-3, Practorcare, NDS and many others allow selecting of nutrients to be printed so that the nutrition counselor can present only calories and fat grams if that is the educational goal. Displaying selected values in graphic form would help clinical dietitians individualize instruction.

References

1. Recommended Dietary Allowances, 10th Edition, Washington, DC: National Academy Press, 1989.

2. Nutrition and Your Health: Dietary Guidelines for Americans, Third Edition, 1990, US Department of

Agriculture, US Department of Health and Human Services, Washington, DC: Government Printing Office.

3. Khoo, Chor San, Industry perspectives on food labeling, Proceedings of the Fourteenth National Nutrient Databank Conference, Ithaca, NY: CBORD, 1989.

4. Stumbo, Phyllis J. Nutrient Calculation Systems for Educators and Clinicians II. Evaluating Data Bases for Microcomputer Programs, Proceedings of the Fifteenth National Nutrient Databank Conference, Ithaca, NY: CBORD, 1991


Figure 1. Graphic display from NAS2 (#10) showing percent of RDA provided by sample menu.

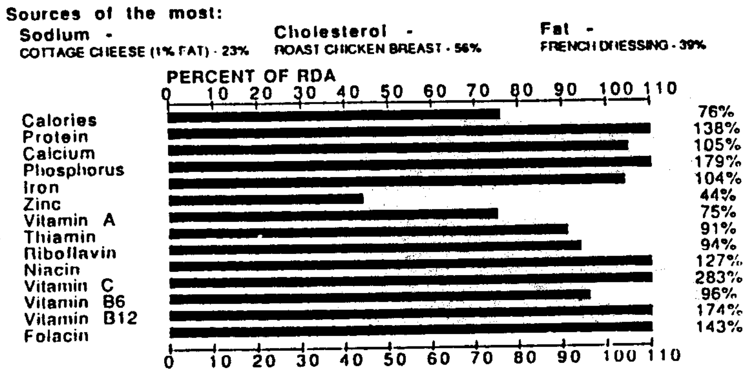



Figure 2. Graphic display from CBORD (#1) showing percent of RDA provided by sample menu.

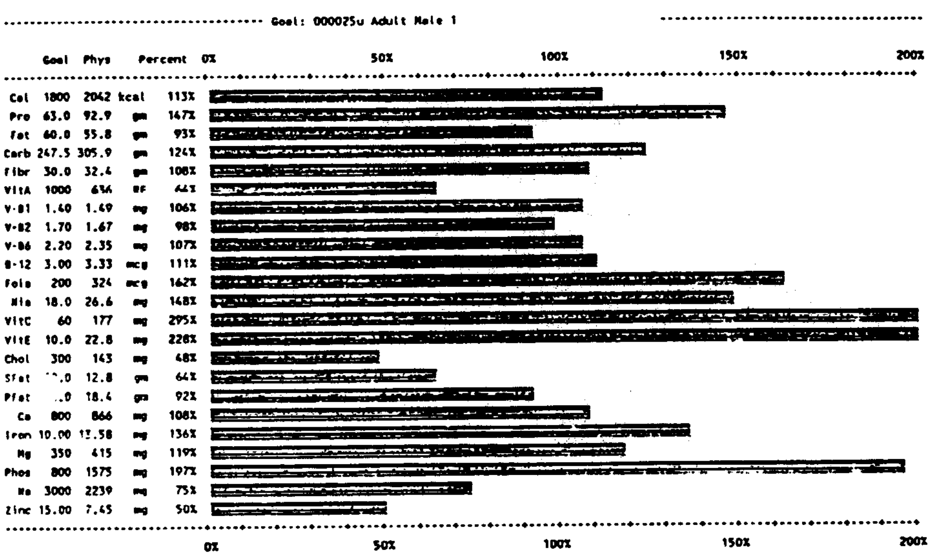



Figure 3. Cholesterol intake recorded at 3 month intervals compared to the study goal for cholesterol intake.

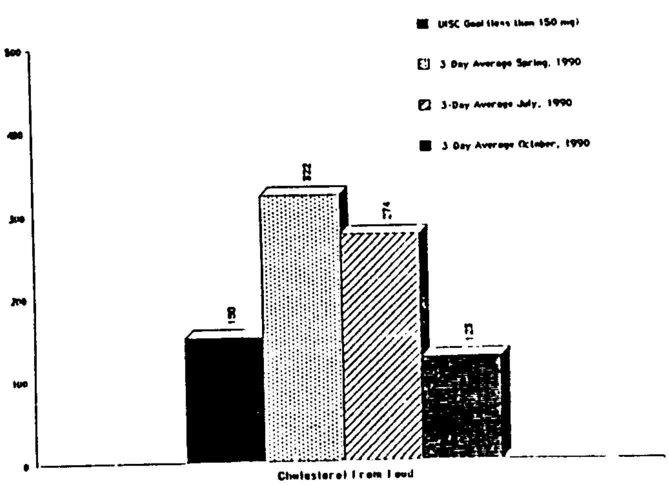
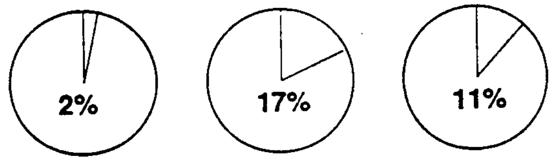


Figure 4. - Pie Charts from Frozen Dinner Label




Of Calories from Saturated Fat

Of Maximum Daily Cholesterol

Of Maximum Daily Sodium

*(6% of Calories from Total Fat)*

National Cholesterol Education Program Recommendations*
- Less than 10% of calories from saturated fat.
- An average of 30% of calories from total fat.
- Less than 300 mg of cholesterol per day.

National Research Council**
- Reccomendation of a maximum sodium intake, of 2,400 mg per day.

* The National Cholesterol Education Program, which is coordinated by the National Heart, Lung and Blood Institute, dows not endorse any particular product.

Figure 5. Pie charts showing calorie distribution of sample menu from NAS2 (#10)

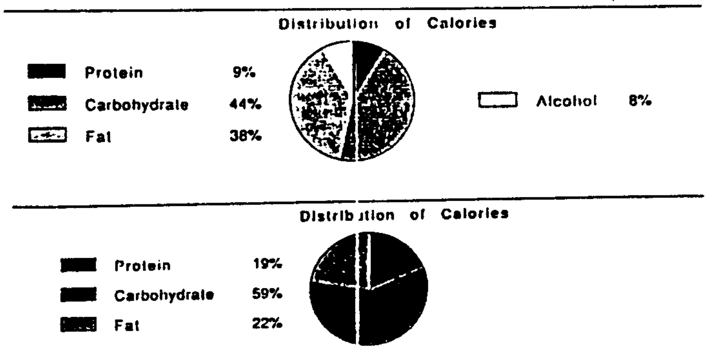



Figure 6. Graphic display from Rite Byte (#12) showing percent of RDA and other nutrient goals provided by sample menu.

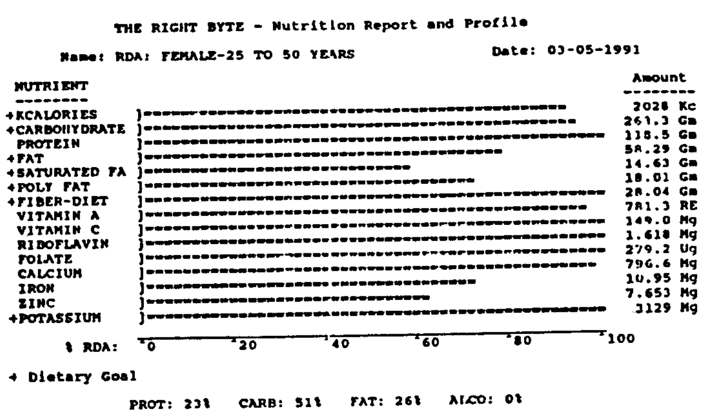



Figure 7. Graphic display from DINE (#5) describing macronutrient composition of sample menu.

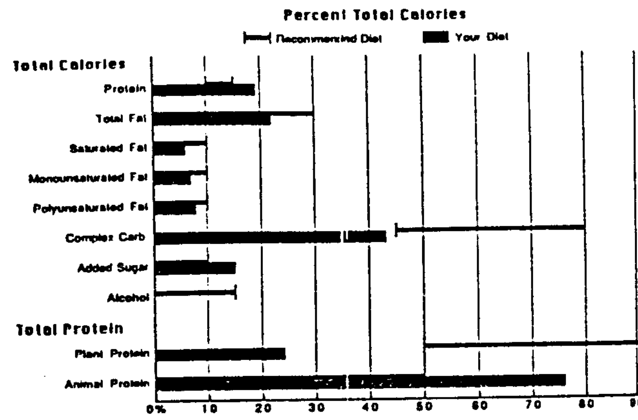



Figure 8. Graphic display from Nutri-tally (#9) showing percent of RDA and other nutrient goals provided by sample diet.

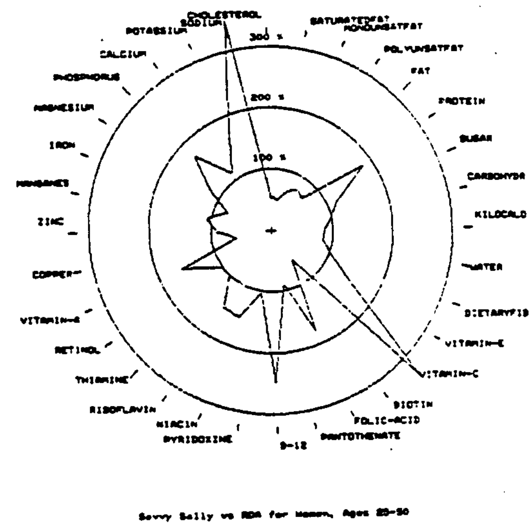
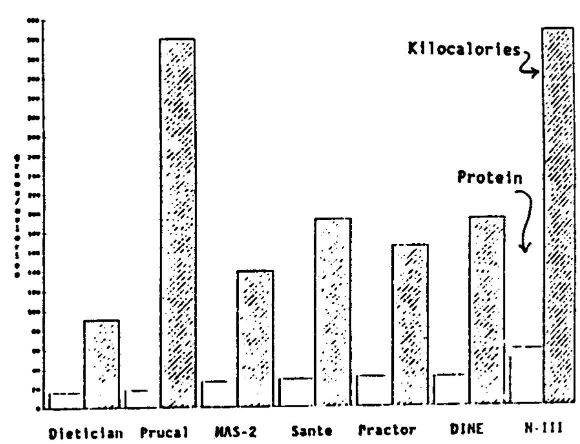



Figure 9. Protein and calories from chicken breast serving.



# Nutrient Variability and Reliability: What to put in a food table?

**William M. Rand+**
**Jean A.T. Pennington***

This presentation addresses the question of nutrient variability from the point of view of what data should be in a food composition table. The starting point is that the user needs information on both placement and spread of each nutrient in each food, and thus the fundamental question is: Given a set of data consisting of replicate measurements of a single nutrient measured in different samples of a single type of food, how should these data be summarized in a few useful statistics?

It needs to be recognized that there are many different reasons why such a set of replicate measurements will not all be identical. Differences will arise because the samples differ genetically and in the environmental conditions under which they developed; after harvest or slaughter, variability is introduced by differences in processing, preparation, and preservation; the samples themselves will differ in how they are prepared for analysis and how they are actually analyzed; and finally, the "food" entry itself may include different items. However, from the perspective of the user of a given food composition table looking at the entry for a particular food, the source of variability is irrelevant. Of interest is only the fact that these "replicates" represent a distribution of what the consumer may consume, and the food table users needs guidance on that point.

Currently each major food composition database has its own conventions for the data it includes, and each presents its data slightly differently. In general, data are included which indicate the nutrient levels most likely to be encountered (the mean, median or mode -- the middle of the distribution), how variable the nutrient is (the standard deviation, the standard error, the range), and how many separate or independent observations these summary statistics are based on. From these statistics, the user estimates typical consumption (from measures of the middle of the distribution) and risks of toxicity and undernutrition (from measures of variability).

Given a set of data, there are a number of well-defined statistics that permit us to estimate where the middle and extremes of a variable are likely to be; however, to interpret anything but percentiles interpolated directly from the data (e.g., median and 5th and 95th percentiles), it is necessary to assume a specific probability distribution (such as normal or lognormal) for each nutrient in each food. This is a strong assumption which allows us to work more confidently with small data sets. The aim of this preliminary study is to explore the basis of making distributional assumptions, and the implications of these assumptions in terms of data presentation in food composition databases. We examined data from nutrient analyses of replicate measurements of more than 200 different foods collected in the Food and Drug Administration's (FDA's) Total Diet Study from 1982 to 1989. We present here our results for two nutrients, sodium and calcium.

The usual distributions that are assumed by workers using food composition data are the normal and the lognormal, and, at one level, skewness distinguishes between these. If a distribution of nutrient content is not skewed, it can often be assumed to be normally distributed, and if the distribution of the log of the nutrient is not skewed, it is assumed to be lognormally distributed. (Throughout, we work with skewness that is normalized, so that a value smaller than -1.96 or larger than +1.96 suggests 0.05 significance.)

Our first observation, illustrated in Figure 1, is that a full range of distributional shapes exists. Note that those that are not skewed could be normal, those skewed positively could be lognormally distributed and those skewed negatively could be neither normal nor lognormal. These are not definitive results since they were selected to illustrate this point from a large number of different foods. These particular histograms may well represent the extremes of inherent sampling variability; however, they do show that categorical assumption of normality or lognormality, for at least sodium and calcium, does not rest on a firm basis.

We next hypothesize that shape of nutrient distribution is related to type of food, and that similar foods will have similar distributions. Figure 2 shows skewness plotted against the food groupings that are used in FDA's Total Diet Study (arbitrarily assigned the first 31 integers). This shows that, contrary to our expectations, food grouping is not closely related to the skewness of the distribution.

Finally, we compared, for each food, the skewness of the nutrient (to see if it might be normally distributed) with the skewness of the log transform of that nutrient (to see if it might be lognormally distributed). Figure 3 shows the results for the two nutrients, sodium and calcium. Based on the skewness, in almost half of the foods looked at, the nutrient could be **either** normal **or** lognormal while in about a sixth of the foods these nutrients could be **neither** normal **nor** lognormal.

This preliminary data exploration leads us to the following conclusions and recommendations:

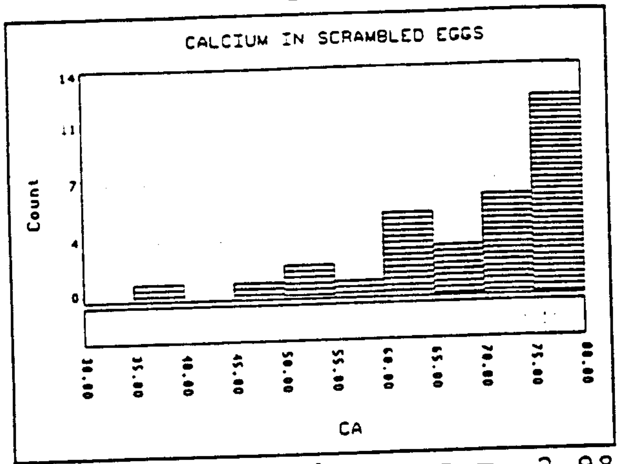
1. We cannot specify with any certainty the distribution of any nutrient in any food; therefore, food composition databases should, as a default, list interpolation estimates of the 5th, 50th (median) and 95th percentiles.

2. The distribution of nutrient in a food appears uncorrelated with the type of food; therefore, it is likely that distribution shape may result from sources of variability other than genetics. Data currently exist which contain information on this and should be analyzed.

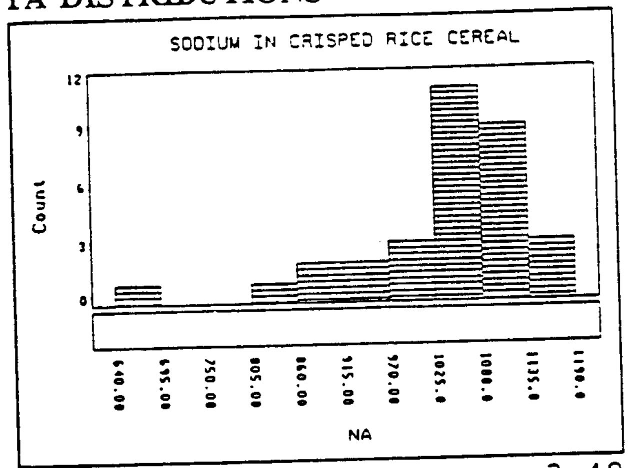
+ William Rand
Tufts University
Boston, Massachusetts

* Jean Pennington
Food and Drug Administration
Washington, District of Columbia

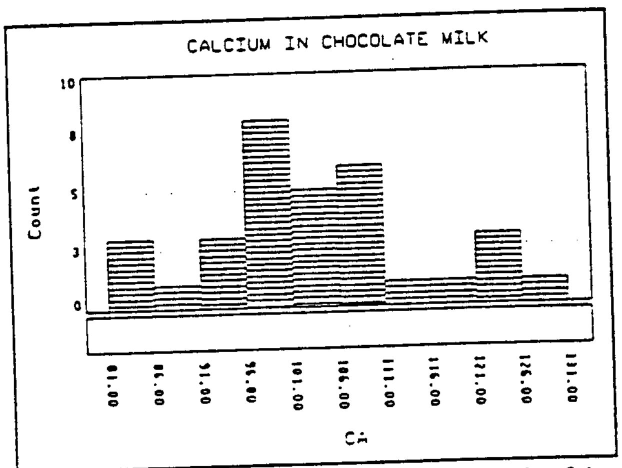
# FIGURE 1: SELECTED DATA DISTRIBUTIONS



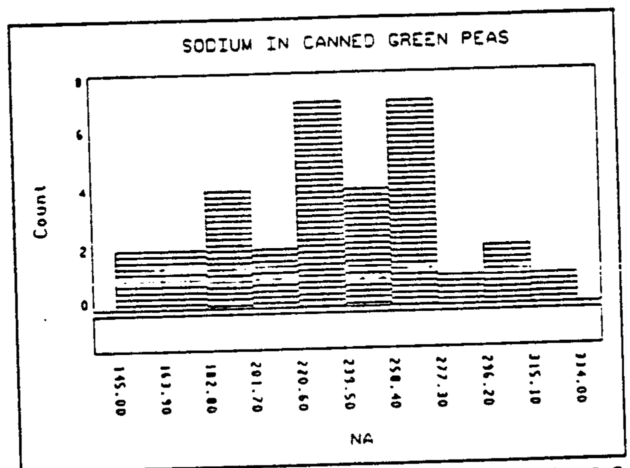

Skewness = -2.98



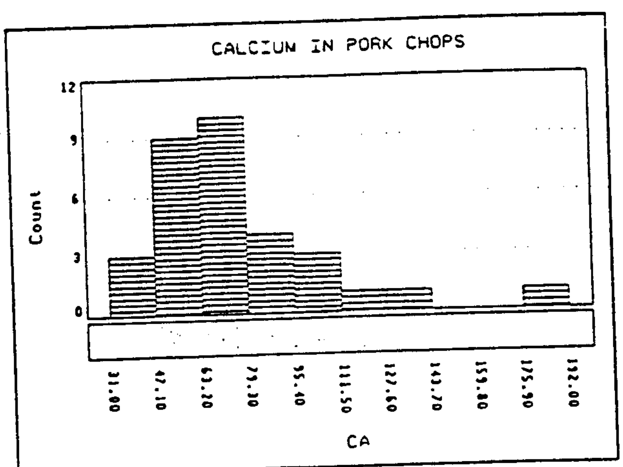

Skewness = -3.49



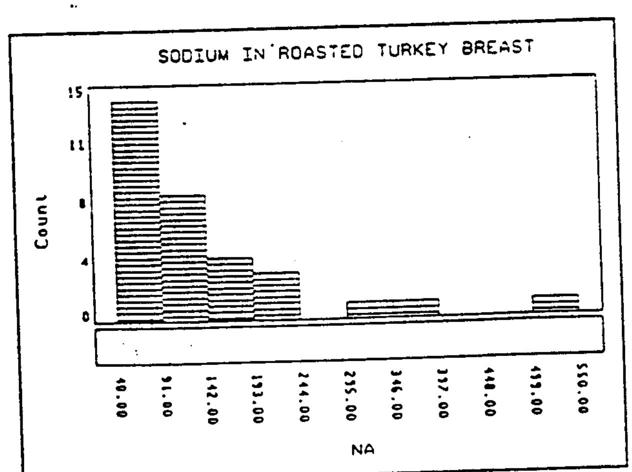

Skewness = 0.64




Skewness = -0.10

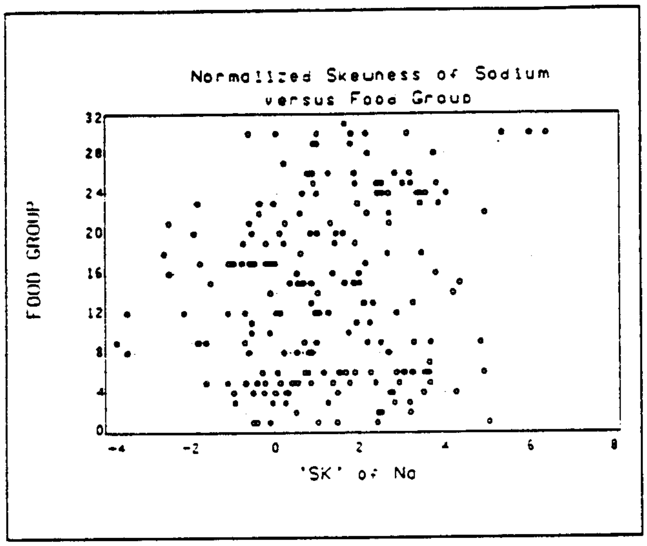



Skewness = 3.51




Skewness = 4.18

FIGURE 2: SKEWNESS VERSUS FOOD GROUP



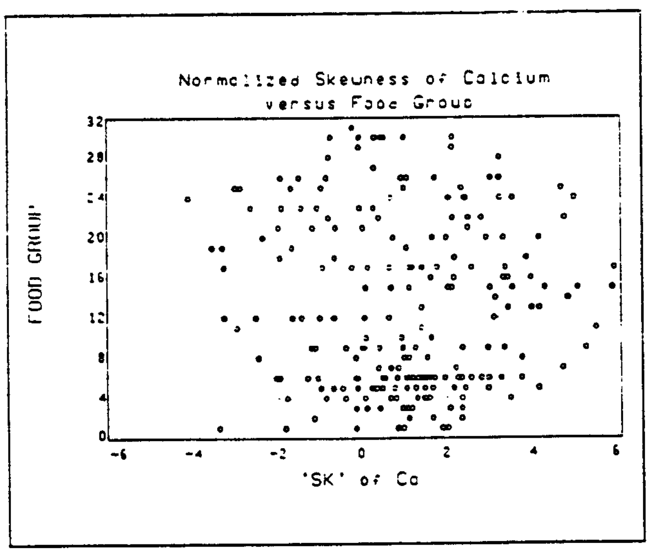
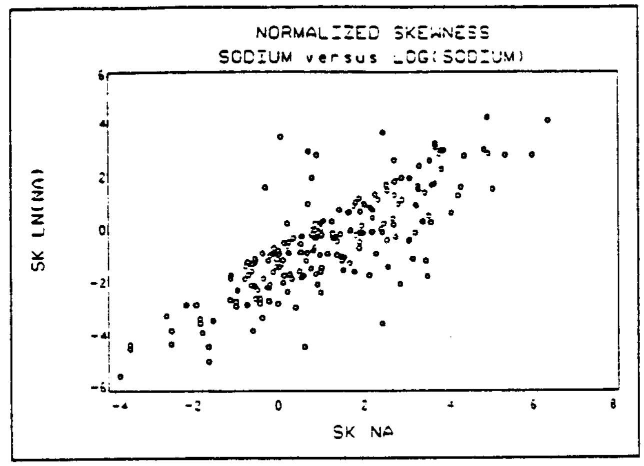
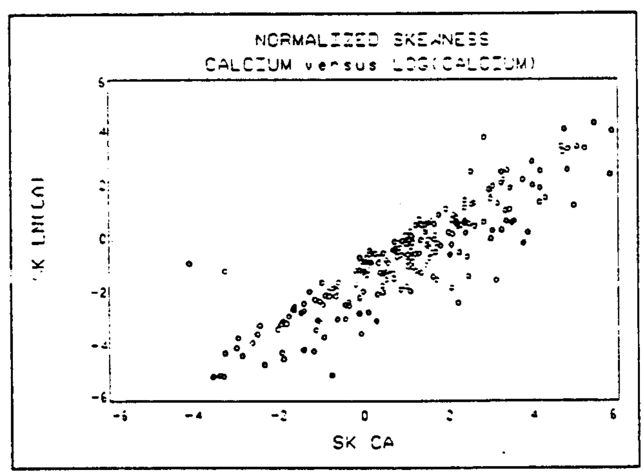




FIGURE 3: SKEWNESS VERSUS LOG SKEWNESS




| Sodium Distribution | No. of Foods | Percent |
|---|---|---|
| Either Normal or Log-Normal | 95 | 47.0% |
| Normal but not Log-Normal | 34 | 16.8 |
| Log-Normal but not Normal | 45 | 22.3 |
| Neither Normal nor Log-Normal | 28 | 13.9 |
| Total Number of Foods Examined | 202 | |




| Calcium Distribution | No. of Foods | Percent |
|---|---|---|
| Either Normal or Log-Normal | 102 | 45.7% |
| Normal but not Log-Normal | 36 | 16.1 |
| Log-Normal but not Normal | 50 | 22.4 |
| Neither Normal nor Log-Normal | 35 | 15.7 |
| Total Number of Foods Examined | 223 | |

# Variability of Estimates of Nutrients in Foods

William Horwitz
Food and Drug Administration
Washington, District of Columbia

The variability of the reported concentrations of nutrients in foods arises from differences in (a) sampling (which includes inherent growing, processing, and distribution factors as well as the physical removal of representative portions); (b) methodological factors (use of different methods of analysis, particularly for method-specific analytes such as "fibers"); (c) operational performance by analysts and laboratories; and (d) interpretation of results (statistical analysis, removal of outliers, categorizing). A major factor for improving the reports of nutrient analysis would be incorporating quality control for these variable factors into all investigations of food composition and including in the final manuscript statements of the quality control specifications used and the extent to which they had been met.

Although scientists realize that there is a certain amount of variability in their measurements, very few have any idea as to its magnitude. Regulatory agencies in particular are sensitive to this variability because, before taking expensive legal action, they wish to be very sure that they are right. Consequently, these organizations will repeat their measurements and even take additional samples to verify their initial findings, all on a case-by-case basis. As a minimum, analytical results in the Food and Drug Administration (FDA) are checked by a second analyst before legal action is approved; in some cases a second laboratory is also used. These replicate measurements are rarely identical, and through long experience regulatory officials may assign "working tolerances" of how much allowance they will make for these differences. In many cases, the party responsible for the goods under investigation can be expected to report results which are at variance with those of the regulatory agency.

Initially, much of this variability was ascribed to methodology. Certainly different methods of analysis could be expected to give somewhat different results. To remove this potential source of variability, regulatory chemists over the past century evolved a verification system which required that any method used for enforcement must be validated by an interlaboratory study to demonstrate the performance characteristics of the method (1). Such an interlaboratory study is conducted by submitting a set of identical, homogenous test samples to a group of typical laboratories for analysis as unknowns. The final results must show acceptably low variability. "Acceptably low" in this context means that the results are usually close to each other and that this pattern is consistent with historical performance. Historical performance has usually been based upon the experience of individual laboratories and is usually summarized by a statement such as, "the results from proximate analysis should agree within plus or minus a few percent and the results on trace nutrients should agree within $\pm 10\%$." With sufficient experience, laboratories set up control limits on a statistical basis so that, for example, 5% of the values generated by the measuring system can lie outside of the boundaries formed by plus and minus 2 standard deviations from the mean. When chemical results do not appear to follow a normal distribution or when considerably more than 5% of the data points transgress these boundaries, an investigation of the source of the excessive variability is warranted.

At the request of the Food and Agriculture Organization (FAO)/World Health Organization (WHO) Food Standards Program (Codex Alimentarius) Committee on Food Labeling, we in FDA have been reviewing the precision of the methods of analysis available for declaring on labels the nutrient content of food. We realized immediately that analytical measurements contribute only one part of the apparent overall variability of the nutrient content of food. The other important components, and in many cases the major components, are (a) the sample, defined to include natural, intrinsic variation in the food, and (b) analytical performance, defined to include method, laboratory, and analyst operations, as shown in Figure 1.

**Figure 1. The major sources of variability of nutrient data and extreme examples from each source.**

**Source of Variability**

| Sampling | Methods | Performance |
|---|---|---|
| Arises from the heterogeneity of the commodity | Different methods respond to different components differently | Analysts perform differently from Training Experience Environment |
| Commodity and Lot Specific | Method Specific | Laboratory and Analyst Specific |
| Extreme Examples: (low): Soybean oil | Iron | Weight |
| (High): Frozen Dinner | Dieatary Fiber | Infrared spectroscopy |
| | Usually confounded (inseparable) | |

If analytical error is less than 1/3 the sampling error, lowering analytical error is of little importance in reducing total erro -- Youden

*Sampling*

The term "sample" will be used here in a very broad sense as covering all variability arising from agricultural, processing, and distribution factors, as well as including the selection and removal of representative portions from representative lots. This broad usage must be carefully distinguished from the narrow attribute of mere removal of representative portions of a food from a specific lot. We will use terms such as a "sampling variability" when referring to the broad concept and "sample error" or "simple sampling" when referring to the narrow concept. The indiscriminate selection of the results from reports of analysis of samples purporting to broadly represent a food is undoubtedly one of the main contributors to variability in database records.

There is not much that can be done with sampling. To obtain a truly representative sample of any important commodity would require designing a sampling plan whose implementation would consume more resources than would be available for an entire database project. Undertaking a suitable sampling plan for a commodity that represents only a fraction of a percent of the food intake of a population is too unimportant to even consider. Consequently many compromises must be made at this point, the chief one of which is that you must take whatever data happen to be available. The best that can be done, in the absence of experimental data, is to estimate (guess), on the basis of experience, the variability that is likely to be encountered from the inherent differences arising from agricultural, processing, and distribution sources. Then the data acceptance process must be managed to ensure that the input data are maintained in statistical control as determined on the basis of the historical input. Data outside of the established control limits require an investigation as to the occurrence of possible mistakes or blunders, which are very difficult to discover after the fact. One helpful factor is that nature does not tolerate gross discrepancies; experience and common sense teach what values do not belong to a category. Therefore, removal of out-of-line data is not the difficult decision-making process that so often occurs in handling data when no information is available to assist in indicating the likely values.

However, even if sampling is confined to the narrow definition of removal of a representative portion from a lot, many papers purporting to supply nutrition data give little information regarding this type of "simple sampling"; therefore, it is impossible to provide a realistic estimate of the sampling variability either with time or with the characterizing, descriptive parameters of the food.

In connection with the preparation of databases, it should also be kept in mind that if a food was originally analyzed for a non-nutritional purpose, the analytical results may be biased with respect to nutritional purposes. For example, regulatory authorities may collect samples because they suspect contamination from chemicals or from filth. They will collect nonrepresentative ("focused") samples intended to contain the contaminant rather than seek a typical sample of the food.

When a sample reaches the laboratory, the first task of the chemist is to reduce that sample both in bulk and in fineness to a manageable size. As a result of this operation, a homogenous mass should be produced which is not expected to introduce any further sampling errors. Consequently, any variability exhibited by this analytical sample is assumed to be entirely the result of analytical operations.

*Analytical Variability*

The major contributors to analytical variability are the inherent biases of the method and the errors introduced by the analyst in the application of the method to a particular material.

Method performance is usually estimated *a priori* by organizations such as the Association of Official Analytical Chemists (AOAC), which compile manuals of approved methods. These societies perform studies by distributing homogenous materials to laboratories to be analyzed as unknowns by the method being tested to determine the fundamental variability among laboratories when the method is used by typical chemists. The variability found in these studies is expected to reflect variability exhibited in actual practice, although it is well known that the variability shown during a method-performance trial is usually less than the variability found in actual practice.

A century-old record exists in the *Journal of the Association of Official Analytical Chemists* of method-performance trials of approved procedures for the analysis of foods. During the first half-century, the data were merely tabulated and the reviewers were left to draw their own conclusions. During the past quarter-century, statistics were used to summarize and to analyze the data, largely as a result of the stimulating lectures of the late Dr. William J. Youden, then of the National Bureau of Standards (2). However, because of the presence of outliers in all measurements and the lack of a standard procedure for examining them, a uniform procedure for reporting interlaboratory analytical data did not exist. Fortunately, in 1987, the International Union of Pure and Applied Chemistry (IUPAC) produced a protocol, designated "IUPAC-1987," for the design, conduct, and interpretation of interlaboratory method-performance (collaborative) studies (3). This protocol, which includes a standard outlier removal procedure, has been accepted by the AOAC and numerous other methods-standardizing organizations in the food field such as the International Dairy Federation (IDF), International Association for Cereal Science and Technology, American Oil Chemists' Society, and International Commission for Uniform Methods for Sugar Analysis.

We first applied this harmonized IUPAC protocol to many of the method-performance studies conducted by the AOAC and the IDF on milk products (4). These methods, particularly for solids, fat, and protein, have had the benefit of over a century of fine-tuning. For the purpose of comparing variability (precision) across several orders of magnitude, we have to use the relative standard deviation, RSD, which is simply the ordinary standard deviation divided by the mean and then placed on a percentage scale by multiplying by 100. The ordinary standard deviation is expressed in the same units as the mean and consequently varies directly with the mean; RSD is dimensionless and is independent of the units, %, g/100 g, decimal fraction, mg/L, etc., and is often independent of concentration.

Figure 2 shows the distribution of the among-laboratories relative standard deviations ($RSD_R$) as a function of concentration for all of the 673 individual data sets in the milk products database. The concentration, expressed as a decimal fraction (where 1% = 0.01) is plotted on the x-axis on a logarithmic scale so that all of the data can be presented in a single figure. Also, it is given as a negative function to "open up the zero". If the conventional type of graph with 0 at the origin of the x-axis was used, many of the low concentration data sets would accumulate near the y-axis, and we would be unable to see their relationship to decreasing concentrations. Distinct clusters are seen at certain concentrations which correspond to the 3% fat and 3% protein of milk (-log 0.03 = 1.6), the 30% fat of cream (-log 0.3 = 0.5), the 40% moisture of cheese (-log 0.4 = 0.4), and the 0.1% phosphorus in milk (-log 0.001 = 3). It is readily seen that the y-dimension of the clusters (variability as relative standard deviation) increases with decreasing concentration (which is equivalent to an increasing [negative] logarithm).

The dotted lower line in Figure 2 is a grand summary of the $RSD_R$ values for over 6000 interlaboratory data sets that we have examined for all types of analytes from aluminum to Zoalene, by methods which range from classical gravimetric analysis to modern mass spectrometry, at concentration levels from pure materials (100%; C = 1.0) to residues and contaminants at a fraction of a part per billion ($C = 10^{-9}$), in solids, liquids, and gases, and in matrices (commodities) which include air, blood, cosmetics, drugs, feeds, fertilizers, ores, pesticides, tissues, and water, as well as foods. Table 1 gives some $RSD_R$ values at useful

**Table 1 - Typical among-laboratories relative standard deviations ($RSD_R$) as a function of concentration expressed as a decimal fraction and conventionally.**

| Concentration | | $RSD_R$ |
|---|---|---|
| Fractional | Conventional | (%) |
| 1.00 | 100% | 2 |
| 0.01 | 1% | 4 |
| 0.0001 | 0.01% | 8 |
| 0.000001 ($10^{-6}$) | 1 ppm | 16 |
| $10^{-8}$ | 10 ppb | 32 |
| $10^{-9}$ | 1 ppb | 45 |

concentration levels, taken from the summary curve. The upper curve of Figure 2 is twice the summary curve values and represents what we consider the upper limit of acceptable precision for interlaboratory studies. This upper curve is a ceiling, based upon our review of the $RSD_R$ values of methods that have been accepted over the past century as approved methodology by the AOAC and other organizations.

All the points and curves on the figures and in the discussion should be considered as "fuzzy"; i.e., all averages, values, and parameters are surrounded by confidence intervals whose width depends upon the desired level of confidence of being right, or its converse, the acceptable level of risk of being wrong. An occasional value beyond the limits can be tolerated; having many values lying beyond the boundaries, however, calls for an investigation as to the cause.

An important aspect about this general curve is that it represents a first approximation, which is useful in the absence of overriding information. Figure 2 shows that $RSD_R$ values of method-performance studies for milk products are somewhat better than those for run-of-the-mill data sets. When we look at the corresponding values from proficiency studies, in which the analyst is not restricted as to the method to be used, we often find that the $RSD_R$ values for the data sets are somewhat worse than those for the method-performance studies. Another important point is that $RSD_R$ refers to the among-laboratories precision. All chemists think that they can do better than the performance shown by the curve, and they can. Within-laboratory precision values, designated as $RSD_r$, are roughly one-half to two-thirds of the among-laboratories values. It is always the other laboratories that are inflating the $RSD_R$ values!

Similar figures applying to more restricted groups have also been prepared from method performance studies of the major nutrients in food -- protein, carbohydrates, and fat; for the supplementary analytes needed to obtain carbohydrates by difference -- ash, moisture, and fibers (5); and for the major mineral elements in food -- calcium, magnesium, phosphorus, potassium, and sodium (6). These graphs are shown in subsequent figures.

Figure 3 shows the precision of a very well-behaved analyte, protein, as a function of C. Each of the 208 $RSD_R$ values for protein from the food database (5) is represented by an upper-case "P," and each of the 201 $RSD_R$ values from the milk products database (4) is represented by a lower-case "m." All the milk values and many of the food values are below the typical lower curve. All but one of the values are below the upper limit. The variability of protein analyses in the concentration range of 1 to 100% can be characterized by an $RSD_R$ of 2%, with most values within the range of 1-3%. An occasional value near 4% in a series is acceptable within the 1-100% concentration range, but having many values above 4% is not acceptable. Within-laboratory variabilities as measured by $RSD_r$ are about one-half of the $RSD_R$ values. This range of 1-3% for $RSD_R$ of protein analyses is further confirmed by the results from laboratory proficiency programs that have been conducted by the American Association of Cereal Chemists and by the American Oil Chemists's Society since the 1920s.

These results may be contrasted with the corresponding among-laboratories results for carbohydrates, fat, and fiber shown in Figure 4. The typical and limit lines are the same, of course, as in Figure 3, but in Figure 4 many individual values are above the upper limit line. Some of the $RSD_R$ values approximate 100%. Many points have $RSD_R$ values above 25% and many of these points are at concentration levels below about 10%. In Figure 5 we show just the fat results, omitting 23 points with $RSD_R$ values above 25% at the concentration level below about 3% fat, "F," in nonmilk foods. Note that all the values from milk products, "f," in Figure 5 are below the "typical" line. The 23 omitted values cannot be considered as being in statistical control, whereas the values for fat in milk products are in excellent control. An important contributor to the high variability of the results for this analyte, fat, as well as those for moisture, ash, and fiber-related components, is the use of too small a test portion (colloquially but incorrectly called "sample weight"). If the test method is applied in such a manner that less than 50 mg of volatiles (moisture) or residue (ash, fat, and fiber components) must be weighed in the final measurement, high variability cannot be avoided. Some of the other conclusions of this review were that low-fiber foods containing less than about 5% fiber cannot be analyzed reliably, regardless of "improved" methodology; that because of disagreements on definitions (accuracy), it is impossible to obtain more reliable methods of analysis; that within-laboratory precision cannot predict among-laboratories precision; and that it is inappropriate to apply methods for fiber to products with very low "fiber" content (flour and rice at 0.5% and starches at 0.1%) and to fluid matrices such as milk and eggs.

Another group for analytes for which data are available are the major elements, shown in Figure 6 (6). Typically most values for these 5 major elements -- calcium, magnesium, phosphorus, potassium, and sodium -- are below the upper limit, but an appreciable

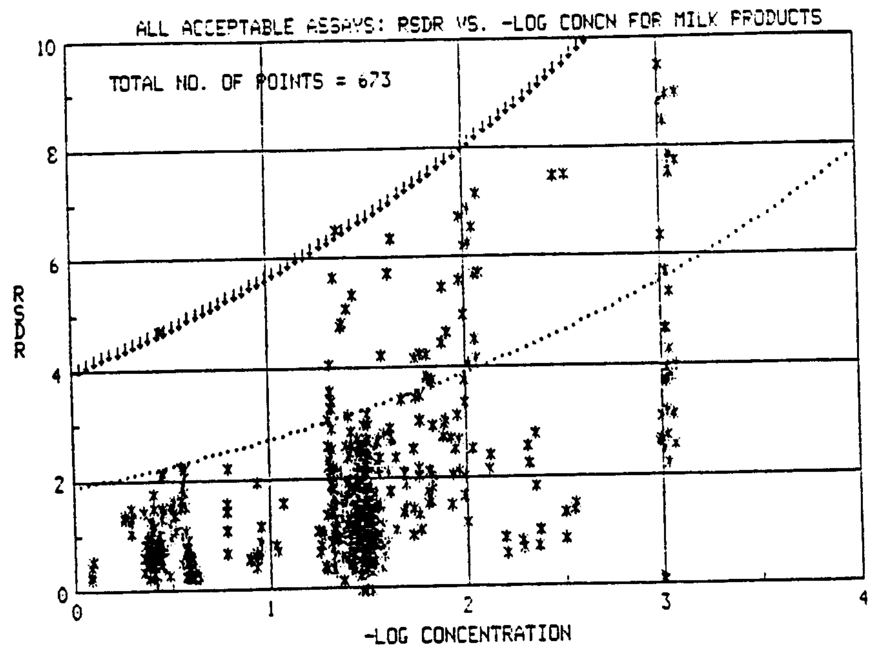



Figure 2. The among-laboratories relative standard deviations, $RSD_R$, for all 673 data sets of analytes (moisture/solids, ash, carbohydrates (by difference), fat, fiber-related, protein, individual sugars, and individual major elements) in milk products as a function of $-\log_{10}C$, where C is expressed as a decimal fraction. The lower line is represented by the equation, $RSD_R = 2^{(1 - 0.5 \log_{10}C)} = 2C^{-0.1505}$; the upper line is twice this curve and is considered the upper empirical acceptable limit for all analytes, independent of analyte, matrix, and method. (Figure 4 from ref. 4.)

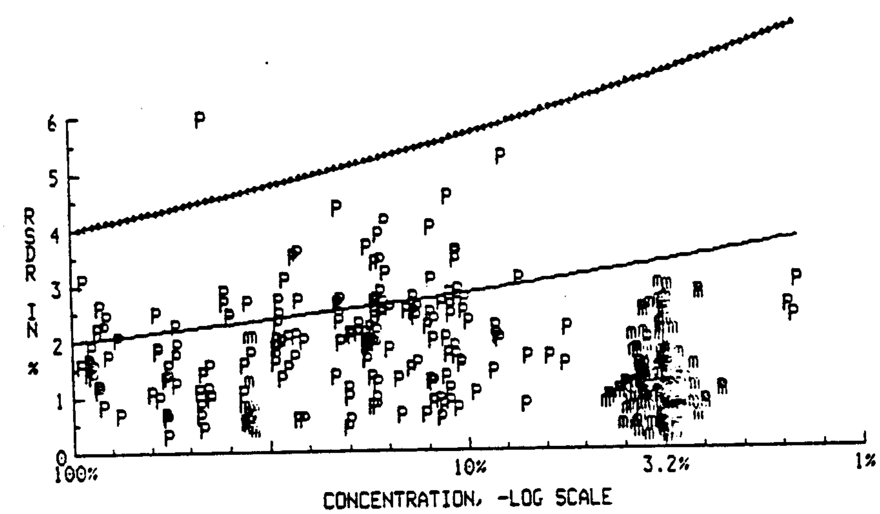



Figure 3. The among-laboratories relative standard deviations, $RSD_R$, for 208 data sets for protein in non-milk foods, P, and from 201 data sets for protein in milk products, m, plotted as a function of -log C as in Figure 2.

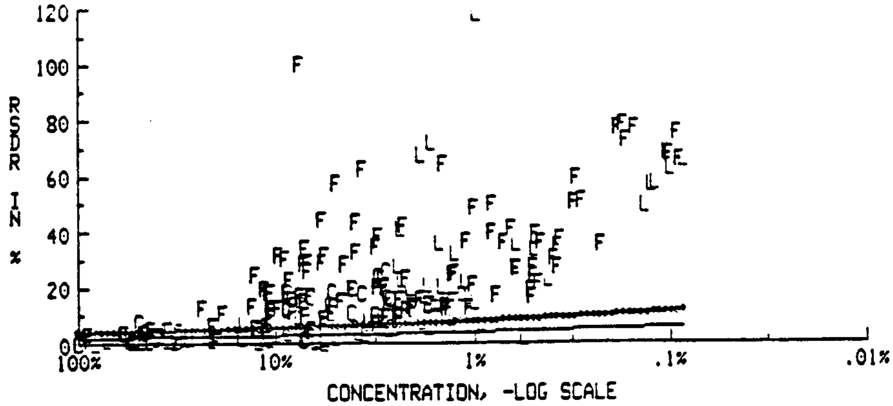



Figure 4. The among-laboratories relative standard deviations, $RSD_R$, for 107 data sets for fiber-related analytes, F; 60, carbohydrates, C; and 112, fat (lipids), L; all from foods, plotted as a function of -log C as in Figure 2.

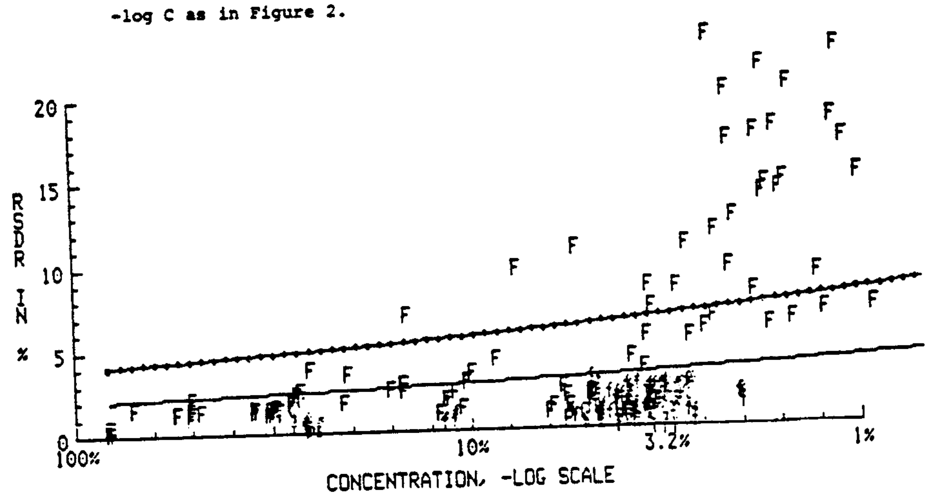



Figure 5. The among-laboratories relative standard deviations, $RSD_R$, for 89 data sets for fat in non-milk foods, F, and from 214 data sets for fat in milk products, f, plotted as a function of -log C as in Figure 2.

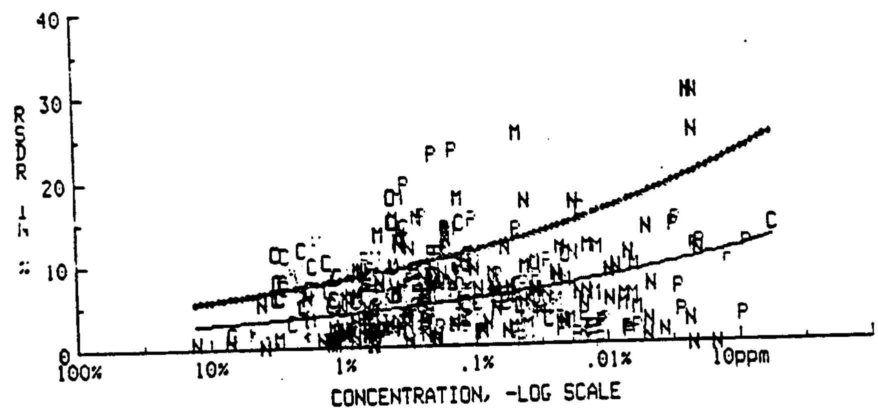



Figure 6. The among-laboratories relative standard deviations, $RSD_R$, for 89 data sets for calcium in foods, C; 48 magnesium, M; 125 phosphorus, P; 91 potassium, K; and 129 sodium, N; plotted as a function of -log C as in Figure 2.

fraction of values exceed the upper boundary of acceptability. This excessive variability is not confined to any particular element or to high or low levels. We interpret these data to indicate that food chemists have not paid much attention to the necessity for quality control of their work but have depended upon familiarity with their procedures to produce the "correct" results. Furthermore, only now are certified reference materials becoming available that will permit food chemists to calibrate their performance against "true" values. With regulatory agencies now expecting the application of "good laboratory practices," with more consideration being paid to laboratory and analyst performance, and with laboratory accreditation being instituted in many countries, particularly within the European common market, much greater attention must be paid to the question of reliability of analytical results.

We are now extending our review of the available interlaboratory studies of methods for the minor and trace elements such as copper, manganese, selenium, and zinc; to vitamins; and to other compounds of nutritional interest such as cholesterol and amino acids.

*Analyst Performance*

Analyst performance is usually difficult to isolate because it is automatically tied up with methodology. In those cases where an attempt is made to separate performance from methods by having the same analyst use different methods, almost invariably the analyst shows much lower variability with the method that is in routine use in the laboratory. Most interlaboratory proficiency studies permit the analyst to use any method. Some studies are so designed that the method effect may be isolated. In most of these cases, when the study shows good control, the variability shown can be represented by the curves previously discussed. Rarely is performance any better than shown in the historical curves; frequently performance is considerably worse. The American Association of Cereal Chemists circulates proficiency materials to be analyzed for the enrichment ingredients niacin and thiamin, for which performance is considerably better than would be expected from the general precision curve. This is easily explained by the fact that well-standardized methods are used routinely by trained technicians who have all the operations on the homogenous test samples from uniform, standardized commodities under excellent control. For those laboratories that conduct these analyses sporadically, however, it can be expected that their performance would approach the general curves.

Experience with examination of analytical results indicates that out-of-control operation usually results from two main sources: (a) clerical errors in recording and transcribing numbers and in performing calculations and (b) improperly prepared standard solutions used for the calibration curves. Clerical and mathematical blunders can be controlled to a large extent by automating these operations. Incorrectly prepared reagents can be discovered through routine quality control procedures such as comparing a current calibration curve with historical data, and through the use of certified reference materials or even the routine use of "house" historical standards.

Routinely participating in professional proficiency studies and taking corrective action when problems are discovered is one means for maintaining the optimum performance of any laboratory. The best way is by randomly using blind in-house check standards and displaying the results as control charts like those often seen now in hospital laboratories. Such quality control operations should be a part of the normal operating budget and management control of laboratory functions.

*Interpretation of Results*

Compilers of databases are familiar with the pitfalls that can come from having to guess at the meaning of various aspects of the results of chemical analysis that authors have neglected to mention. These aspects may be the simple omission of factors used in the conversion of nitrogen to protein or a fundamental failure to indicate the steps that were taken to ensure the validity of the results. The need for validation of results is particularly important in food analysis because so many of the methods of analysis are method-specific (empirical). As indicated previously, in many cases the differences may not be very great; but in the case of dietary fiber such differences are the cause of endless polemics, even though the absolute differences are not very significant from the point of view of nutritional labeling. An absolute difference of 1% in dietary fiber in a product that contains 10% relative difference, will change the carbohydrates (by difference) by only 1% or 4 calories/100 g. This difference is of the same order of magnitude as the $RSD_R$ (10%) shown by the interlaboratory studies of methods for fiber-related analytes at the 10% concentration level, so reasonable allowances must be made for this variability.

The term "reasonable allowances" deserves further discussion. Regulatory officials are well aware of the variability produced by inherent processing factors and by chemical analysis. They make due allowance for the inherent random variability from these sources. But the type of variability they make allowances for is the 2-sided variability that appears both above and below the target (labeled) amount. A series of analytical results that appear consistently on the low side of a label declaration typically is not a result of random variability but is rather (a) a symptom of a real deficiency of analyte, or (b) a bias in the performance of the analysis, which bias can be isolated by the use of the regulatory method (1) or by use of a certified reference material, if available. The fact that the value is within the confidence interval of the analytical and sampling error is not a defense against a deviation in the declaration of a nutrient. There is a clear danger inherent in the use of nutrient data obtained by the uncritical pooling of numbers found in the literature through the potential introduction of unsuspected biases. The resulting average or range may have little relationship to reality. The literature of food analysis is notorious for this type of error.

*Discussion*

It is difficult to generalize with respect to the relative importance of the factors contributing to the variability of results that may enter a database. Unless individual portions from a lot (technically known as increments) have been analyzed in the original work, and their deviations carried over to a database, the error ascribable even to "simple" sampling cannot be estimated. Very few investigations have been performed with the objective of determining this type of sample error, with all other variables -- method, operator, storage, etc. -- held constant. Most such investigations have been interested in other phenomena such as variety, geography, agricultural conditions, and production and distribution variables. In many investigations of nutrient content as a function of food production, the extent of sampling and analytical errors has not even beenconsidered.

Food chemists have never thought it important to assess the variability inherent in sampling for various reasons. Most commodities are purchased on an "as is" basis or by a contract. If the value is based on a specification, the laboratories and the methods of analysis are often specified. An arbitration procedure for settling economic disputes is usually a built-in requirement. In the case of promulgation of food standards by the FDA during the 1940s and 1950s, representative nationwide samples of many foods were obtained. The sampling and analytical variabilities were incorporated into the final standard by specifying a limit close to the minimum (or maximum, as the case required) found in commercial channels.

Even knowledge of "simple sampling" variability is important because often the method or the analyst is blamed for poor performance when actually the initial laboratory sample or the prepared analytical sample is responsible for the 'obvious variability. Unrepresentative sampling can introduce a large uncontrollable error, for which no allowance can be made *ex post facto*.

Methodology is probably not too important a contributor to the variability seen in nutrient databases except in a few very well-known cases such as dietary fiber. Even when different methods such as Kjeldahl nitrogen or Dumas nitrogen are applied to the same food, the difference in the final result is only a matter of a few tenths of a percent. The same is probably true when different times and temperatures are used in the determination of total solids, moisture, and ash. A small but significant difference may be apparent if chromatography is used for the separation and determination of individual sugars and the results are compared to the previously reported results for the same bu combined group of "total reducing sugars, before and after inversion." The recent collaborative studies conducted by Tanner and Barnett (7) verify the applicability of the current AOAC methods for nutrients in foods in general to milk-based infant formulae. Practically all of the results from these studies bracket th general curve shown previously. The final phase o these studies, which are now under review for publication, includes apparently satisfactory method perfor mance data for nutrients in foods for which there ha been no official methods, such asiodine and vitamin K One interesting and important aspect of the use of th general precision curve is the finding that the method for vitamin D are already quite acceptable. Th precision data (among-laboratories), $RSD_R$ = 20-40% in the literature do not reflect poor laboratory perfor mance or poor methods but merely that this vitamin being determined at levels of less than 1 ppb, for whic an $RSD_R$ of 45% is acceptable on historical grounds

Probably the most important single factor respon sible for variability in food constituent databases is th absence of quality control of the analytical work. Eve if quality control existed in the reporting laborator statements to this effect are frequently missing in th final manuscripts. Only within the last few years hav certified reference materials become available to pra ticing food chemists to provide an absolute referen point for their analyses (8) for the control of accura of those analytes that are not method-dependent.

One potential improvement in the situation, making food chemists aware of the importance quality control, is the application of the "expert sy tem" concept developed by the U.S. Department of Agriculture (USDA), Beltsville Human Nutrition Research Center, to published nutrient data (9). This system evaluates the published reports and data and assigns a quantitative rating scale based on the number of laboratory samples reported, the validity of the analytical methods, the handling and documentation of the laboratory and analytical samples, the sampling plan, and the extent of analytical quality control, all important aspects of good analytical practices (10). The summation of the quality factors for each item results in an overall 3-factor confidence code (A (best), B, and C), indicating the relative degree of confidence the user can have in a grand mean value for the analyte in a food. The system has been applied to selenium (11) and copper (12) in food. In the case of copper, only 14% of the confidence codes for 218 foods for which reports had been examined rate A, 24% rated B, and 62% rated C (limited confidence due to limited data quantity and/or quality). In general, the large number of C ratings was an indication of the paucity of data. Although it is possible that quality control may have been performed as part ofthe published investigations but not mentioned in the manuscripts, it is more likely that little attention was given to this critical requirement. In the case of some important foods, the widely disparate literature values required implementation of confirmatory analyses by the USDA laboratory. The scheme now requires modification to take into consideration the current availability of certified reference materials as an alternative to the original use of multiple laboratories to evaluate analytical bias.

These findings suggest that not much confidence can be placed in data that reside in the literature unaccompanied by documentation that they were produced under controlled conditions. This situation is not confined to inorganic elements, for which certified reference materials of reasonable similarity to the foods of interest are available as a check on the correctness of the results. Holden and Davis (13) describe their experience in selecting an analytical contractor for the analytical phase of a nationwide survey to update the USDA nutrient data for eggs. The results from blind analysis of National Institute of Standards and Technology (NIST) Standard Reference Material (SRM) 1563-2, Cholesterol and Fat-Soluble Vitamins in Coconut Oil, were used in part to select a technically competent contractor. The certified values was 624 µg/g with an acceptable range of 601-674 µg/g; only 2 of the 5 commercial and university laboratories submitted results within the limits of acceptability set by NIST. The reported results of analyses arranged in increasing order were 218, 287, 607, 643, and 866. Replication of results within a laboratory is not an acceptable quality control technique. The triplicate determinations in all laboratories indicated good precision. Only the availability of a reference value salvaged this phase of the study. Even the mean of these results, 524 µg/g, was outside the acceptable range, and the relative standard deviation of the 5 values was an unacceptable 50%. Government con-

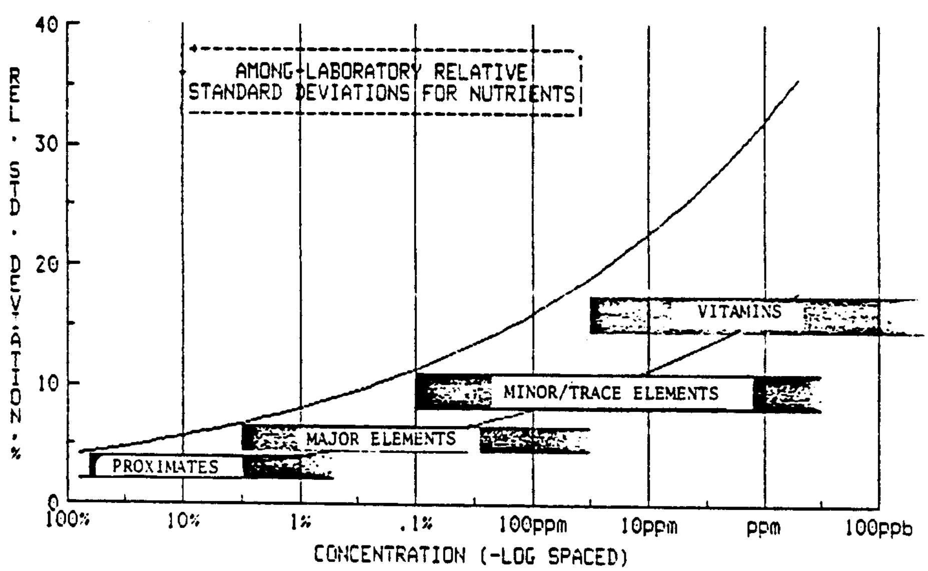



Figure 7. The summary graph of the among-laboratories relative standard deviations ($RSD_R$ in %) expected for the major nutrient groups in relation to their concentration level as present in any food. The curves through the four rectangles and near the upper left corners are the same curves appearing in Figures 2-6.

tracting officers can relate similar experiences. In one case a request for proposals for pesticide residue analyses elicited interest from several dozen potential contractors. The number dropped to a half dozen when it was disclosed that the award would be made in part on the basis of results from actual blind analyses of the commodities of interest.

### *Conclusions*

As a result of our review of the results of approximately 6000 interlaboratory studies conducted under fairly well-controlled conditions, we have constructed Figure 7, a summary of the $RSD_R$s to be expected for the results of analyses for the various nutrients by a group of laboratories. The values do not include allowance for variability inherent in the commodity itself or from growing, processing, and sampling factors. (If these factors are also present, the standard deviations must be squared to obtain the variances, the variances should be added, and the square root of the sum should be taken to obtain the final "total" standard deviation. This standard deviation is transformed to an RSD by dividing by the mean and then multiplying by 100. A better value can be calculated with a scientific calculator (with an exponent key) from the general formula given at the bottom of Table 1, by inserting a specific concentration expressed as a decimal fraction.) The results are considered only as an approximate historical descriptive summary of the data of numerous studies. Any individual study can deviate considerably, by as much as a factor of 2 in either direction, and still be within the acceptable confidence interval.

The production of a database, however, requires the existence of absolute standards against which the values for specific analytes can be measured. In the case of proximate analysis, the analytical results are method-specific with no systemic error. No NIST standards existed during the data accumulation phases. Formulation for these types of analytes of reference materials that would remain stable over a reasonable period of laboratory storage would be difficult, although dried eggs and dried milks kept refrigerated and in a moisture- and oxygen-free atmosphere approach reference material requirements. But these foods are not useful for the important fiber-related analytes. Incorporation of a section in published papers describing the quality control efforts on an equal basis with the classical sections on materials, methods, results, discussion, conclusion, and references should become a necessary part of good manuscript preparation practices.

Table 2 - **Some typical and maximum among-laboratories acceptable relative standard deviations** ($RSD^R$) to be expected from historical analytical variability of nutritionally important analytes. Sampling and fabrication variability, if present, must be added vectorially (as variances).

| Nutrient | Concentration | | | $RSD_R$ | |
|---|---|---|---|---|---|
| | Mean | Unit | Range | Typical | Maximum |
| Proximates | 10 | % | 100-0.5 | 2 | 5 |
| Major Elements | 0.1 | % | 5-0.005 | 5 | 10 |
| Minor/Trace Elements | 10 | ppm | 1000-0.5 | 10 | 20 |
| Vitamins | 1000 | ppb | 50-5000 | 15 | 35 |

### *References*


(1)Food and Drug Administration (1990) *Code of Federal Regulations* Title 21, Section 2.19, U.S. Government Printing Office, Washington, DC 20402.

(2)Youden, W. J. (1967) *J. Assoc. Off. Anal. Chem.* 50, 1007-1013. *Statistical Manual of the AOAC,* Association of the Official Analytical Chemists, Arlington, VA 22201.

(3)Horwitz, W. (1988) *Pure Appl. Chem.* 60, 855-864; (1989) *J. Assoc. Off. Anal. Chem.* (1989) 72, 694-704.

(4)Peeler, J. T., Horwitz, W., and Albert, R. (1989) *J. Assoc. Off. Anal. Chem.* 72, 784-806.

(5)Horwitz, W., Albert, R., Deutsch, M. J., and Thompson, J. N. (1990) *J. Assoc. Off. Anal. Chem.* 73, 661-680.

(6)Horwitz, W., Albert, R., Deutsch, M. J., and Thompson, J. N. Under review for submission to *J. Assoc. Off. Anal. Chem.*

(7)Tanner, J. T. and Barnett, S. A. (1985) *J. Assoc. Off. Anal. Chem.* 68, 514-522; (1986) 69, 777-785; additional manuscript under review.

(8)Rasberry, S. (May 1991) *Am. Lab.* 23 (8), 59-60.

(9)Bigwood, D. W., Heller, S. R., Wolf, W. R., Schubert, A., and Holden, J. M. (1987) *Anal. Chim. Acta* 200, 411-419.

(10)Horwitz, W. (1978) *Anal. Chem.* 50(6), 521A-524A.

(11)Schubert, A., Holden, J. M., and Wolf, W. R. (1987) *J. Am. Diet. Assoc.* 87, 285-299.

(12)Lurie, D. G., Holden, J. M., Schubert, A., Wolf, W. R., and Miller-Ihli, N. J. (1989) *J. Food Comp. Anal.* 2, 298-3[illegible]6.

(13)Holden, J. M. and Davis, C.S. (1990) *Fresenius J. Anal. Chem.* 338, 476-478.

# Comments on Variability in Food Composition Data

**Kent K. Stewart**
**VPI & SU**
**Blacksburg, Virginia**

The variability of food composition data arises from:

* the real differences in the foods,
* the variability inherent in the process of the selection of the samples from the total food supply,
* the variability due to the lack of complete homogenization of the food samples,
* the variability due to the lack of reproducible extraction of the analyte from the food samples,
* the variability due to the lack of reproducible analytic separation, detection and measurement of the analytes,
* the variability due to the lack of reproducible computation of the results of the assays.

Thus the total variance of the food composition values is a function of the variances listed above. From the users' point of view what is wanted is a measure of the real differences in the composition of foods. Thus it is crucial that the variances due to sampling techniques, sample preparation techniques, and the analytic techniques and computations be known so that the real variances of food compositions can be determined.

Some way must be found of determining the variances due to the measurement processes. At present, there is insufficient information to determine the most common sources of variances in food composition data. My guess is that the source of the variances of the data varies from food to food and from nutrient to nutrient.

# Are we all Speaking the Same Language - A Reprise

**Grace J. Petot**
**Case Western Reserve University**
**Cleveland, Ohio**

At the Eighth National Nutrient Data Bank Conference in Minneapolis in 1983, I presented a review of the documentation of nutrient data bases and nutrient analysis systems citations in the Journal of the American Dietetic Association and the American Journal of Clinical Nutrition published from January 1983 through June 1983. (Petot, 1983). At that time, I stated that "nutrient data base systems continue to proliferate in attempting to meet the needs", and that such "activities are the basis for our common interests but we persist in speaking different languages as we separately pursue them". Today, eight years later, the situation has changed with respect to the numbers of nutrient data base systems available, but it has not really changed with respect to the descriptions or documentation of data sources.

The discussion today will focus on "Publishing and citing papers concerning nutrient data". I will summarize a review of documentation of nutrient data systems or data sources in the 1990 issues of the Journal of the American Dietetic Association, The American Journal of Clinical Nutrition and the American Journal of Epidemiology. I hope that this review will raise some questions for consideration and discussion.

Each of the panelists will also identify issues. We probably won't produce any solutions to the problems today, but we offer a forum for communication and, even if we don't all speak the same language, at least we may offer some translations which may be better understood by the creators of nutrient data base systems, the variety of users, and the entire research community which is the ultimate consumer of the data.

## DOCUMENTATION OF COMPUTERIZED NUTRIENT ANALYSIS IN PUBLISHED RESEARCH

In the 1982-83 survey, 42 investigations in which computerized nutrient data bases were used were reported in the Journal of the American Dietetic Association and the American Journal of Clinical Nutrition. It is quite astounding today to realize that no one was using microcomputers in these investigations. The data base systems in use then were all on mainframe computers. At that time, 22 of 42 authors identified the data base system by name (Table 1). Only two of those 22 cited a date or version number of the system, and 11 of 22 did not cite any references of documentation for the data used. Twenty of 42 did not identify the systems used.

In 1990, seven years later, the situation was quite different. All 1990 issues of the Journal of the American Dietetic Association, Volume 90, the American Journal of Clinical Nutrition, Volumes 51 and 52, and the American Journal of Epidemiology, Volumes 131 and 132 were surveyed to identify all reported investigations which used dietary intake information for calculated analyses using computerized nutrient data bases or nutrient data base systems (Table 2 and APPENDIX).

These journals were selected because they contain numerous dietary intake reports which seem to be representative of similar reports which may be scattered among many other research journals. There were 119 articles in which computer calculations were made for 1 - 31 nutrients in 24-hour recalls, diet diaries, weighed diet records, weighed diets, and quantified food frequency records. The number of subjects in the reported studies ranged from 1 - 88,837 (Table 3).

Summarizing the documentation of sources of information presented an even greater challenge than it did 8 years ago. In a 12 month period there were almost three times as many reports of computerized dietary intake analyses as there were during the eighteen month period eight years ago.

Sixty reports identified a nutrient data base system by name; twenty-seven of those reported a date or version of the system (Table 4). Thirty-six different United States systems were identified. Nine studies used foreign systems.

Many different methods were used to cite or reference food composition tables and data base systems (Table 5). Eighty three food tables and systems were named in the text or cited in references. Twenty seven data base systems were documented by date of the system or by a version number. For twenty nine computerized data bases, it was stated (usually in the text) that it was developed by the user or at the institution where the work was done. For most of these data bases, references to food composition tables were made. There was such a mixture of these that it was not possible to tally accurately all of the references to data sources. The only consistencies noted were that in the JADA, in 15 of 46 reports, a footnote was used in the text for the name and address of proprietary data base systems; and in the AJCN, in 18 of 66 reports, parentheses were used to enclose the name and address of these data base systems. Otherwise, there seemed to be no consistency or policy with respect to documenting food composition data used for analysis. In Table 6 is a list of data sources mentioned in the articles. Many of them were expressed in different ways; some were mentioned both in the text and with references; others were mentioned only in the text; many different names were used for the same food tables, etc. A list of some examples of expressions used in the text is in Tables 7.1 and 7.2. Note that for many, there was no further documentation.

Editorial policies for most peer reviewed journal reports require scientific documentation for research methods and procedures. The science community should require it of food composition data. The task for us, as ones who are interested in appropriate and accurate uses of food composition information, is to help create some order to the confusion and inconsistency and to assist researchers, grant reviewers and editors with guidelines.

Petot, G.J. Commonalities and differences: Are we all speaking the same language? In: Tobelman, R., Editor. Proceedings of the Eighth National Nutrient Data Bank Conference. July 25-27, 1983. Minneapolis, MN. U.S. Department of Commerce, NTIS. Springfield, VA

**Table 1 - NDB Systems Used Cited JADA, AJCN 1/82-6/83**

| Sources Documented | Unidentified N = 20 | Identified N = 22 |
|---|---|---|
| Bulletin 72 | 3 | 0 |
| Handbook 8 1963 | 6 | 2 |
| Handbook 456 | 10 | 3 |
| Rev Handbook 8 | 5 | 0 |
| Bowes & Church | 5 | 1 |
| Literature | 6 | 5 |
| Industry | 4 | 2 |
| Laboratory analysis | 0 | 1 |
| Imputed | 1 | 0 |
| Other | 7 | 4 |
| None | 3 | 11* |

* Only 2 w/date or version no.

**Table 2 - Journals Surveyed**

Journal of the American Dietetic Association

Vol 90 Jan - Dec 1990

American Journal of Clinical Nutrition

Vols 52,52 Jan - Dec 1990

Amercan Journal of Epidemiology

Vols 131, 132 Jan-Dec 1990

**Table 3 - Journal Survey Summary 46 JADA 66 AJCN 7 AJE**

119 journal articles

1-31 nutrients

1-88,837 subjects

54 studies diaries

12 studies weighed food records

10 studies quantified food frequencies

9 sutdies 24-hr recalls

9 studies weighed menus or meals

8 studies QFF + diaries

7 studies recalls + diaries

10 studies miscellaneous methods

**Table 4 - Nutrient Data Base Systems Named**

60 studies identified a system by name

27 studies cited date/version of system

36 different U.S. systems identified

9 studies used foreign systems

**Table 5 - Method for Citing In Text or References**

| Documentation | Number |
|---|---|
| NDB or NDB System Named | 83 |
| NDB/NDB system date or version | 27 |
| Developed by user | 29 |
| Cited only in references | 23 |
| No name, source or reference | 4 |
| Footnore in text JADA (n=46) | 15 |
| () in text AJCN (n=66) | 18 |

**Table 6 - Nutrient Data Sources Cited In Text or References**

Handbook 8 1963
Revised Handbook 8
Standard Reference NDB/Tapes
Handbook 456
Bulletin 72
Bowes & Church
Pennington
Literature
Manufacturers
Laboratory Analysis
Calucuated Recipes
Foreign food tables (by name)
Journal references (not nutrient data)
Unpublished manuscript

**Table 7.1 - Nutrient Data Sources as Mentioned in Text I**

- "based on USDA nutrient data"
- "computer analysis"
- "continuously updated with USDA data"
  "based on UDA and manufacturers nutrient composition data and other sources"
  "system based on USDA and other publishedreferences"
  "3500 brand name and generic...plus decisions on 45 foods added"
- "plus user added items"
  "developed at" (name of institution)
- "1500 fooods to which additional items could be added"

- no further documentation

**Table 7.2 - Nutrient Data Sources as Mentioned in Text II**

- "based on Handbook 8 and supplemental data from the research literature and manufacturers"
- "calcuated from standard food tables" (Bull 72 1986, Pennington 1980)
  "nutrient data base at the CRC" (Hdbk 456, Bowe & Church 1980)
- "computer file of more than 3000 food items"
  "3500 brand name and generic...plus decisions on 45 foods added"
- "using an updated version of a computer data base"
- " the results were presentaed as mean daily intakes of..."
- "multiple references available on request"

## APPENDIX
## REFERENCES CITED IN LITERATURE SURVEY


1. Buzzard, M.I., Chlebowski, R.T., Jeffery, R.W. et al. Diet intervention methods to reduce fat intake: Nutrient and food group composition of self-selected low-fat diets. J Am Dietet A 90:42 1990
2. Briley, ME., Montgomery, D.H. and Blewett, J. Dietary intakes of police department employees in a wellness program. J Am Dietet A 90:65 1990
3. Shah, M., Laing, B., Savre, S.G. et al. Hypertension prevention trial (HPT): Food pattern changes resulting from intervention on sodium, potassium, and energy intake. J Am Dietet A 90:69 1990
4. Hou, S.C. and Khan, M.A. Comparative assessment of sodium intake from meals and snacks consumed by selected American and British students. J Am Dietet A 90:97 1990
5. Harkey, A.L. Clinical effects of mandibular fixation on cholesterol intake and serum lipids: a case study. J Am Dietet A 90:102 1990
6. Kristal, A.R., and Hololy, A.L. Patterns of dietary behavior associated with selecting diets low in fat: Reliability and validity of a behavioral approach to dietary assessment. J Am Dietet A 90:214 1990
7. Demark-Wahnefried, W., Bowering, J. and Cohen, P.S. Reduced serum cholesterol with dietary change using fat-modified and oat bran supplemented diets. J Am Dietet A 90:223 1990
8. Larkin, F. A., Basiotis, P.P., Riddick, H.S. et al. Dietary patterns of women smokers and non-smokers. J Am Dietet A 90:230 1990
9. Hickson, J.F., Johnson, T.E., Lee, W. and Sidor, R.J. Nutrition and the precontest preparations of a male bodybuilder. J Am Dietet A 90:264 1990
10. Chima, C.S., Miller-Kovach, K., Zeller,M. et al. Lipid managemant clinic: Dietary intervention for patients with hypercholesterolemia. J Am Diet A 90:272 1990
11. Pennington, J.A.T. and Wilson, D.B. Daily intakes of nine nutritional elements: Analyzed vs. calculated values. J Am Dietet A 90:375 1990
12. VanHorn, L.V., Gernhofer, N., Moag-Stahlber, A. et al. Dietary assessment in children using electronic methods: Telephones and tape recorders. J Am Dietet A 90:412 1990
13. Wilson-Goodman, V., Taylor, M.L., Mueller, D. and Palmer, J. Factors affecting the dietary habits of adolescents with cystic fibrosis. J Am Dietet A 90:429 1990
14. Engle, A., Hebert, J.R. and Reddy, B.S. Relationships between food consumption and dietary intake among healthy volunteers and implications for meeting dietary goals. J Am Dietet A 90:526 1990
15. Alford, B.B., Blankenship, A.C. and Hagen, R.D. The effects of variations in carbohydrate, protein, and fat content of the diet upon weight loss, blood values, and nutrient intake of adult obese women. J Am Dietet A 90:534 1990
16. Suitor, C.W., Gardner, J.D. and Feldstein, M.L. Characteristics of diet among a culturally diverse group of low-income pregnant women. J Am Dietet A 90:543 1990
17. Greene, G.W., Achterberg, C., Crunmbaugh, J. and Soper, J. Dietary intake and dieting practices of bulimic and non-bulimic female college students. J Am dietet A 90:576 1990
18. Fong, A.K.H. and Kretsch, M.J. Nutrition evaluation scale system reduces time and labor in recording quantitative dietary intake. J Am Dietet A 90:664 1990
19. Mcintosh. W.A., Kubena, K.S., Walker, J. et al. The relationship between beliefs about nutrition and dietary practices of the elderly. J Amer Dietet A 90:671 1990

20. Morrow, S.R. and Mona, L.K. Effect of gastric balloons on nutrient intake and weight loss in obese subjects. J Am Dietet A 90:717 1990
21. Gorbach, S.L., Morrill-LaBrode, A., Woods, M.N. et al. Changes in food patterns during a low-fat dietary intervention in women. J Am Dietet A 90:803 1990
22. Potosky, A.L. and Hartman, A.M. The apparent validity of diet questionnaires is influenced by number of diet-record days used for comparison. J Am Dietet A 90:810 1990
23. Treiber, F.A., Leonard, S.B., Frank, G. et al. Dietary assessment instruments for preschool children: Reliability of parental responses to the 24-hour recall and a food frequency questionnaire. J Am Dietet A 90:814 1990
24. Sanjur, D., Garcia, A., Aguilar, R. et al. Dietary patterns and nutrient intakes of toddlers from low-income families in Denver, Colorado. J Am Dietet A 90:823 1990
25. Manore, M.M., Vaughan, L.A. and Lehman, W.R. Contribution of various food groups to dietary vitamin B-6 intake in free-living, low- income elderly persons. J Am Dietet A 90:830 1990
26. Mangels, A.R., Moser-Veillon, P.B., Andon, M.GB. and Reynolds, R.D. Vitamin B-6 intakes of lactating women: Analyzed vs calculated values. J Am Dietet A 90:857 1990
27. Berman, M.K., Taylor, M.L. and Freeman, E. Vitamin B-6 in premenstrual syndrome. J Am Dietet A 90:859 1990
28. Kleiner, S,M., Bazarre, R.L. and Litchford, M.D. Metabolic profiles, diet, and health practices of championship male and female bodybuilders. J Am Dietet A 90:962 1990
29. Pike, S., Prati, R.C., Cochran, C. et al. Lack of effect of vitamin B-6 supplementation on the lipoprotein profile of post- menopausal chronic hemodialysis patients. J Am Dietet A 90:968 1990
30. Carrillo, T.E., Gilbride, J.A. and Chan, M.M. Soup kitchen meals: An observation and nutrient analysis. J Am Dietet A 90:989 1990
31. Howat, P.M., Varner, L.M. and Wampold, R.L. The effectiveness of a dental/dietitian team in the assessment of bulimic dental health. J Am Dietet A 90:1099 1990
32. Martinez, J.A., Urbistondo, M.D. and Velasco, J.J. Assessment and implications of the dietary intakes of hospitalized psycho- geriatric patients. J Am Dietet A 90:1111 1990
33. Ashraf, H., Schoeppel, C. and Nelson, J.A. Use of tofu in preschool meals. J Am Dietet A 90:1114 1990
34. Teufel, N. I. and Dufour, D.L. Patterns of food use and nutrient intake of obese and non-obese Hualapai Indian women of Arizona. J Am Dietet A 90:1229 1990
35. McCorkindale, C., Dybevik, K., Coulston, A.M. and Sucher, K.P. Nutritional status of HIV-infected patients during the early disease stages. J Am Dietet A 90:1236 1990
36. Luder, E., Ceysens-Okada, E., Koren-Roth, A. and Martinez-Weber, C. Health and nutrition survey in a group of urban homeless adults. J Am Dietet A 90:1387 1990
37. Witschi, J.C., Capper, A.L. and Ellison, R.C. Sources of fat, fatty acids and cholesterol in the diets of adolescents. J Am Dietet A 90:1429 1990
38. Bergman, E.A., Boyungs, J.C. and Erickson, M.L. Comparison of a food frequency questionnaire and a 3-day diet record. J Am Dietet A 90:1431 1990
39. Gropper, S.A.S., Anderson, K., Landing, W.M. and Acosta, P.B. Dietary selenium intakes and plasma selenioum concentrations of formula-fed and cow's milk-fed infants. J Am Dietet A 90:1547 1990
40. Tsui, J.C. and Nordstrom, J.W. Folate status of adolescents: Effects of folic acid supplementation. J Am Dietet A 90:1551 1990
41. Nieman, D.C., Onasch, L.M. and Lee, J.W. The effects of moderate exercise training on nutrient intake in mildly obese women. J Am Dietet A 90:1557 1990
42. Stanek, K, Abbott, D. and Cramer, S. Diet quality and the eating environment of preschool children. J Am Dietet A 90:1582 1990
43. Davis, M.A., Murphy, S.P., Neuhaus, J.M. and Lein, D. Living arrangements and dietary quality of older U.S. adults. J Am Dietet A 90:1667 1990
44. Brown, J.E., Tharp, T.M., Dahlberg-Luby, E.M. et al. Videotape dietary assessment: Validity, reliability, and comparison of results with 24-hour dietary recalls from elderly women in a retirement home. J Am Dietet A 90:1675 1990
45. Dubois, S. and Boivin, J. Accuracy of telephone dietary recalls in elderly subjects. J Am Dietet A 90:1680 1990
46. Rhodus, M.L. and Brown, J. The association of xerostomia and inadequate intake in older adults. J Am Dietet A 90:1688 1990
47. Kottler, D.P., Tierney, A.R., Brenner, S.K. et al. Preservation of short-term energy balance in clinically stable patients with AIDS. Am J Clin Nutr 51:7 1990
48. Haste, F.M., Brooke, O.G., Anderson, H.R. et al. Nutrient intakes during pregnancy: Observations on the influence of smoking and social class. Am J Clin Nutr 51:29 1990
49. Coulston, A.M., Mandelbaum, D. and Reaven, G.M. Dietary management of nursing home residents with non-insulin-dependent diabetes mellitus. Am J Clin Nutr 51:67 1990
50. Mattes, R.D., Cowart, B.J., Schiavo, M.A. et al. Dietary evaluation of patients with smell and/or taste disorders. Am J Clin Nutr 51:233 1990
51. Forster, J.L., Jeffery, R.W., VanNatta, M. and Pirie, P. Hypertension prevention trial: do 24-h records capture usual eating behavior in a dietary change study? Am J Clin Nut 51:253 1990
52. Pusateri, D.J., Roth, W.T., Ross, J.K. and Shultz, T.D. Dietar and hormonal evaluation of men at different risks for prostat cancer: Plasma and fecal hormone-nutrient interrelation-ships. Am J Clin Nutr 51:371 1990
53. Motil, K.J., Montandon, C.M., Thotahuchery,M. and Garza, C. Dietary protein and nitrogen balance in lactating anc nonlactating women. Am J Clin Nutr 51:378 1990
54. Harris, W .S., Rothrock, D.W., Fanning, A. et al. Fish oils i hypertriglyceridemia: A dose-response study. Am J Cli Nutr 51:399 1990
55. Greger, J.L., Davis, C.D., Sutti, J.W. and Lyle, B.J. Intake serum concentrations, and urinary excretion of manganese b adult males. Am J Clin Nutr 51:457 1990
56. Kretsch, M.J. and Fong, A.K.H. Validation of a new comput erized technique for quantitating individual dietary intake The Nutrition Evaluation Scale System (NESSy) vs th weighed food record. Am J Clin Nutr 51:477 1990
57. Allman, M.A., Stewart. P.M., Tiller, D.J. et al. Energ supplementation and the nutritional status of hemodialysi patients. Am J Clin Nutr 51:558 1990
58. Minghelli, G., Schutz, Y, Charbonniew, A. et al. Twenty four-hour energy expenditure and basal metabolic rate mea sured in a whole-body indirect calorimeter in Gambian mer Am J Clin Nutr 51:563 1990
59. deGroot, L., van Es, J.H., van Raaij, J. et al. Energ metabolism of overweight women 1 mo and 1 y after an 8-w slimming period.Am J Clin Nutr 51:578 1990
60. Worme, J.D., Doubt, T.J., Singh, A. et al. Dietar patterns,gastrointestinal complaints, and nutrition know edge of recreational triathletes. Am J Clin Nutr 51:690 199
61. Frusztajer, M.T., Dhuper, S., Warren, M.P. et al. Nutritio and theincidence of stress fractures in ballet dancers. Am Clin Nutr 51:779 1990

62. Klesges, R.C., Eck, L.H., Isbell, T.R. et al. Smoking status: Effects on the dietary intake, physical activity, and body fat of adult men. Am J Clin Nutr 51:784 1990

63. Barbosa, J.C., Shultz, T.D., Filley, S.J. and Nieman, D.C. The relationship among adiposity, diet, and hormone concentrations in vegetarian and nonvegetarian postmenopausal women. Am J Clin Nutr 51:798 1990

64. Wardlaw, G.M. and Snook, J.T. Effect of diets high in butter, corn oil, or high-oleic acid sunflower oil on serum lipids and apolipoproteins in men. Am J Clin Nutr 51:815 1990

65. Tordoff, M.G. and Alleva, A.M. Effect of drinking soda sweetened with aspartame or high-fructose corn syrup on food intake and body weight. Am J Clin Nutr 51:963 1990

66. Kesaniemi, Y.A., Tarpila, S. and Miettinen, T.A. Low vs high dietary fiber and serum, biliary, and fecal lipids in middle-aged men. Am J Clin Nutr 51:1007 1990

67. Anderson, J.W., Gustafson, N.J., Spencer, D.B., Tietyin, J. and Bryant, C.A. Serum lipid response of hypercholesterolemic men to single and divided doses of canned beans. Am J Clin Nutr 51:1013 1990

68. Childs, M.T., Dorsett, C.S., King, I.B., Ostrander, J.G. and Yamanaka, W.K. Effects of shellfish consumption on lipoproteinsin normolipidemic men. Am J Clin Nutr 51:1020 1990

69. Lukaski, H.C., Hoverson, B.S., Gallagher, S.K., and Bolonchuk, W.W. Physical training and copper, iron, and zinc status of swimmers. Am J Clin Nutr 51:1093 1990

70. Colditz, G.A., Willett, W.C., Stampfer, M.J. et al. Patterns of weight change and their relation to diet in a cohort of healthy women. Am J Clin Nutr 51:1100 1990

71. Jacob, V., LeCarpentier, J.E., Salzano, S. et al. IGF-I, a marker of undernutrition in hemodialysis patients. Am J Clin Nutr 52:39 1990

72. Kenler, H.A., Brolin, R.E. and Cody, R.P. Changes in eating behavior after horizontal gastroplasty and Roux-en-Y gastric bypass. Am J Clin Nutr 52:87 1990

73. Lovelady, C.A., Lonnerdal, B. and Dewey, K.G. Lactation performance of exercising women. Am J Clin Nutr 52:103 1990

74. Moser-Veillon, P.B. and Reynolds, R.D. A longitudinal study of pyridoxine and zinc supplementation of lactating women. Am J Clin Nutr 52:135 1990

75. van Beresteijn, E.C.H., Brussard, J.H. and van Scheik, M. Relationship between the calcium-to-protein ratio in milk and the urinary calcium excretion in healthy adults - a controlled crossover study. Am J Clin Nutr 52:142 1990

76. Reilly, C., Barrett, J.E., Patterson, C.M. et al. Trace element nutrition status and dietary intake of children with phenylketonuria. Am J Clin Nutr 52:159 1990

77. de Castro, J.D. and Orozco, S. Moderate alcohol intake and spontaneous eating patterns of humans: evidence of unregulated supplementation. Am J Clin Nutr 52:246 1990

78. Flaten, H., Hostmark, A.T., Kierulf, P. et al. Fish-oil concentrate: effects on variables related to cardiovascular disease. Am J Clin Nutr 52:300 1990

79. Bandini, L.G., Schoeller, D.A., Cyr, H.N. and Dietz, W.H. Validity of reported energy intake in obese and nonobese adolescents. Am J Clin Nutr 52:421 1990

80. Miller, W.C., Lindeman, A.K., Wallace, J. and Niederpruem, M. Diet composition, energy intake, and exercise in relation to body fat in men and women. Am J Clin Nutr 52:426 1990

81. Snyder, D.K., Underwood, L.E. and Clemmons, D.R. Anabolic effects of growth hormone in obese diet-restricted subjects are dose-dependent. Am J Clin Nutr 52:431 1990

82. Boyd, N.F., C ousins, M., Beaton, M. et al. Quantitative changes in dietary fat intake and serum cholesterol in women: results from a randomized, controlled trial. Am J Clin Nutr 52:470 1990

83. Schafer, L. and Overvad, K. Subcutaneous adipose-tissue fatty acids and Vitamin E in humans: relation to diet and sampling size. Am J Clin Nutr 52:486 1990

84. O'Dea, K., Trainedes, K., Chisholm, K., Leyden, H. and Sinclair, A.J. Cholesterol-lowering effect of a low-fat diet containing lean beef is reversed by the addition of beef fat. Am J Clin Nutr 52:491 1990

85. Anderson, J.W., Spencer, D.B., Hamilton, C.C. et al. Oat-bran cereal lowers serum total and LDL cholesterol in hypercholesterolemic men. Am J Clin Nutr 52:495 1990

86. Fukagawa, N.K., Anderson, J.W., Hageman, G. Young. V.R. and Minaker, K.L. High-carbohyrate, high-fiber diets increase peripheral insulin sensitivity in healthy young and old adults. Am J Clin Nutr 52:524 1990

87. Calvo, E.B. and Gnazzo, N. Prevalence of iron deficiency in children aged 9-24 mo from a large urban area of Argentina. Am J Clin Nutr 52:534 1990

88. Schneeweiss, B., Graninger, W., Stokenhuber, R. et al. Energy metabolism in acute and chronic renal failure. Am J Clin Nutr 52:596 1990

89. Motil, K.J., Montandon, C.M. and Garza, C. Basal and postprandial metabolic rates in lactating and non-lactating women. Am J Clin Nutr 52:610 1990

90. Nieman, D.C., Haig, J.L., Fairchild, K.S. et al. Reducing-diet and exercise-training effects on serum lipids and lipoproteins in mildly obese women. Am J Clin Nutr 52:640 1990

91. Turnbull, W.H., Leeds, A.R. and Edwards, G.D. Effect of mycoprotein on blood lipids. Am J Clin Nutr 52:646 1990

92. Kestin, M., Moss, R., Clifton, P.M. and Nestel, P.J. Comparative effects of three cereal brans on plasma lipids, blood pressure, and glucose metabolism in mildly hypercholesterolemic men. Am J Clin Nutr 52:661 1990

93. Grense, D.A., Bucuvalas, J.C. and Balistreri, W.F. Efficacy of cornstarch therapy in type III glycogen-storage disease. Am J Clin Nutr 52:671 1990

94. Himmelstein, S., Clemens, T.L., Rubin, A.E. and Lindsay, R. Vitamin D supplementation in elderly nursing home residents increases 25(OH)D but nor 1,25(OH)2D. Am J Clin Nutr 52:701 1990

95. Orwell, E.S. and Oviatt, S. Relationship of mineral metabolism and long-term calcium and cholecalciferol supplementation to bloodpressure in normotansive men. Am J Clin Nutr 52:717 1990

96. Barger-Lux, M.J., Heaney, R.P. and Stegman, M.R. Effects of moderate caffeine intake on the calcium economy of premenopausal women. Am J Clin Nutr 52:722 1990

97. Cano, N., Labastie-Coeyrehowig, J., Lacombe, P. et al. Perdialytic parenteral nutrition with lipids and amino acids in malnourished hemodialysis patients. Am J Clin Nutr 52:726 1990

98. Tsunehara, C.H., Leonetti, D.L. and Fujimoto, W.Y. Diet of second generation Japanese-American men with and without non-insulin dependent diabetes. Am J Clin Nutr 52:731 1990

99. Bennett, F.C. and Ingram, D.M. Diet and female sex hormone concentrations: an intervention study for the type of fat consumed. Am J Clin Nutr 52:808 1990

100. Brown, A.J., Roberts, D.C.K., Pritchard, J.E. and Truswell, A.S.A mixed Australian fish diet and fish-oil supplementation: impact on the plasma lipid profile of healthy men. Am J Clin Nutr 52:825 1990

101. Fricker, J., Le Moel, G. and Apfelbaum, M. Obesity and iron

stores in menstruating women. Am J Clin Nutr 52:863 1990
102. Lutz, J. and Tesar, R. Mother-daughter pairs: spinal and femoral bone densities and dietary intakes. Am J Clin Nutr 52:872 1990
103. Matkovic, V., Fontana, D., Tominac, C., Goal, P. and Chesnut, C.H. Factors that influence peak bone mass formation: a study of calcium balance and the inheritance of bone mass in adolescent females. Am J Clin Nutr 52:878 1990
104. Knekt, P., Albanes, D., Seppanen, R. et al. Dietary fat and breast cancer. Am J Clin Nutr 52:903 1990
105. Patischman, N., McCulloch, C.E., Byers, R. et al. Breast cancer and dietary plasma concentrations of carotenoids and vitamin A.Am J Clin Nutr. 52:909 1990
106. Maclure, K.M., Hayes, K.C., Colditz, G.A., Stampfer, M.J. and Willett, W.C.. Dietary predictors of symptom-associated gallstones in middle-aged women. Am J Clin Nutr 52:916 1990
107. Payette, H., Rola-Pleszczynski, M., and Ghadirian, P. Nutrition factors in relation to cellular and regulatory immune variables in a free-living elderly population. Am J Clin Nutr 52:927 1990
108. Kardjati, S., Kusin, J.A., Schofield, W.M. and deWith, C. Energy supplementation in the last trimester of pregnancy in East Java, Indonesia: effect on maternal anthropometry. Am J Clin Nutr 52:987 1990
109. Creed deKanashiro, H., Brown, K.H., Lopez deRomana, G. and Black, R.E. Consumption of food and nutrients by infants in Huascar (Lima), Peru. Am J Clin Nutr 52:995 1990
110. Brown, K.H., Stallings, R.Y., Creed deKanashiro, H., Lopez deRomana, G., and Black, R.E. Effects of common illnesses on infants energy intakes from breast milk and other foods during longitudinal community-based studies in Huascar (Lima), Peru. Am J Clin Nutr 52:1005 1990
111. Bell, L., Hectorn, K.J., Reynolds, H. and Hunninghake, D.B. Cholesterol-lowering effects of soluble-fiber cereals as part of a prudent diet for patients with mild to moderate hypercholesterolemia. Am J Clin Nutr 52:1020 1990
112. Smidt, L.J., Cremin, F.M., Grivetti, and Clifford, A.J. Influence of folate status and polyphenol intake on thiamin status of Irish women. Am J Clin Nutr 52:1077 1990
113. Graham, S., Marshall, J., Haughey, B. et al. Nutritional epidemiology of cancer of the esophagus. Am J Epidem 131:454 1990
114. van't Veer, P., van der Wielen, R.P.J., Kok, F.J. et al. Selenium in diet, blood, and toenails in relation to breast cancer: A case control study. Am J Epidem 131:987 1990
115. Flegel, K.M. and Larkin, F.A. Partitioning macronutrient intake estimates from a food frequency questionnaire. Am J Epidem 131:1046 1990
116. Farrow, D.C. and Davis, S. Diet and the risk of pancreatic cancer in men. Am J Epidem 132:423 1990
117. Zeigler, R.G., Brinton, L.A., Hamman, R.F. et al. Diet and the risk of invasive cervical cancer among white women in the United States. Am J Epidem 132:432 1990
118. Hartman, A.M., Brown, C.C., Palmgren, J. et al. Variabilityin nutrient and food intakes among older middle-aged men. Am J Epidem. 132:999 1990
119. Romieu, I., Stampfer, M.J., Stryker, W.S. et al. Food predictors of plasma beta-carotene and alpha-tocopherol: validation of a food frequency questionnaire. Am J Epidem 131:864 1990

Panel Presentation

# Publishing and Citing Papers Concerning Nutrient Data

**Elaine R. Monsen, Ph.D., R.D.**
**University of Washington, Seattle**
**Seattle, Washington**


In presenting data generated from nutrient databases, I suggest a minimum to be included of: 1) the name of the database software developer or tradename, 2) the copyright year/year of most recent update, 3) the version number, if applicable, 4) a statement describing modification/additions to the database, 5) a statement, if appropriate, to indicate how missing data, missing foods, and other missing data were handled, 6) a statement, if considered appropriate, as to the source of the nutrient content data (e.g. USDA and/or food manufacturers).

Because of the importance of supplying needed information to readers and researchers, I propose that an Ad Hoc committee be established from this Nutrient Database Conference that can address these issues. Suzanne Murphy, who has been the Chairman of the meeting, would be an excellent leader for such a committee and I am sure that many of us here would be willing to serve with her in developing and working out suggestions of guidelines that would be appropriate to use when data from different databases is presented or published.

In addition to the citing of databases, there are five other issues that are critical in research:

1. The accuracy of the database, its completeness and its currency.
2. Accurate coding of foods consumed.
3. Valid and reliable methods of collecting food intake data.
4. Proper sample selection to assure that the study population is properly identified and adequately sampled.
5. High subject response rate, to assure that the subjects represent the sample which, in turn, represents the study population.

Publishing and Citing Papers Concerning Nutrient Data

# Perspectives of a User, Author, and Reader

**Jean H. Hankin, Ph.D., R.D.**
**University of Hawaii**
**Honolulu, Hawaii**

As a nutritionist in an epidemiologic research group, I am concerned about the comparability of dietary data reported in cohort and case-control studies. Epidemiologists seek confirmation of their findings on diet and disease. If the results differ from other investigations concerning the same dietary component and the same disease, nutritionists and epidemiologists look for possible explanations in the methodologies of the studies. One of the first issues to investigate are the similarities and differences in the dietary methodology and the food composition data.

I concur with Grace Petot, the chairperson of this session, about the frequent absence of information on the sources of food composition data in published papers. Recently, Audrey Maretzki, Editor of the Journal of Nutrition Education, sent me copies of eight papers with dietary intake data that were published in this Journal. In my opinion, only three of these were satisfactorily documented. The authors of one article stated that "computerized nutritional analysis of menu items provided information about nutrient content" (1). The reference was to a paper concerning a computerized analysis system. There was no information about the source of the nutrient data. The text of a second paper indicated that the 24-hour recalls were analyzed by Food Processor II (ESHA Research, Salem, OR) (2). There was no further information, no citation in the references, and no date. In a third article, the authors stated that the "unique data-based design and computer software packet are described elsewhere" (3). The references were to a paper by Block and to a computer system packet from the National Cancer Institute. Again, no information concerning the source of the nutrient data was included. Finally, another paper used a Home and Garden Bulletin from 1981 to analyze 1990 data (4). Although these papers did not pertain to diet and disease associations, they indicate that better documentation is needed.

Because of my interest in the association of carotenoids and cancer risk, I'm always interested in the methods used for assessing their intakes. As you know, until recently, data on the carotenoid contents of food items were sparse. Yet, in a paper on vitamin A, beta-carotene, and cancer risk, published in the Journal of the National Cancer Institute in 1987 (5), I found this statement, "Using the U.S. Department of Agriculture tables of food values for standard portion size of each food item, we estimated the average daily dietary intake of vitamin A and beta-carotene for each individual...." Interestingly, none of the nutrient data were published. Rather, the population was divided into thirds, according to low, medium, or high intakes of vitamin A and beta-carotene; these distributions were compared to the incidence rates of selected cancer sites. I suspect that the investigators utilized the vitamin A values of fruits and vegetables and assumed that the vitamin A in these foods could be considered the beta-carotene content of the diet. This, we know, is questionable.

Another paper dealing with the epidemiologic evidence on the role of carotenoids in reducing cancer risk (6) included two statements that I found misleading and inaccurate. The first stated that "present food composition tables reflect the content of beta-carotene, alpha-carotene, beta-cryptoxanthin, lycopene, and possibly several other carotenoids." This is only partially true since some of these carotenoids do not possess vitamin A activity. A footnote indicated that the

AOAC method does not resolve beta-carotene from the other hydrocarbon carotenoids, but that the xanthophylls are separated. This statement suggests that data on the xanthophyll contents of food items are available. However, it is only very recently that data on selected carotenoids have been published. A second misleading comment stated that "current food composition tables list for carotenoid content a composite value which measures a number of hydrocarbon carotenoid species including beta-carotene." Again, there was a footnote, "food composition tables actually list the vitamin A content of food in International Units and Retinol Equivalents, and the carotenoid content must be calculated from these values." There was no information concerning the methods for calculating these values.

In my opinion, there are indeed some acceptable ways to estimate and report carotene intakes. A few examples from the literature are shown below:

1. Report total vitamin A, and separately, vitamin A from fruits and vegetables (Byers et al, 1987) (7).

2. Compute vitamin A activity consumed in form of retinol precursors (Hinds et al, 1984) (8).

3. Use FAO/WHO percentage distribution of vitamin A into retinol, beta-carotene, and other carotenoids for different food groups, and convert to micrograms of each component (Le Marchand et al, 1989)(9).

4. Use McCance and Widdowson's food composition data for estimating total retinol and carotenes (Paul and Southgate, 1978) (10).

5. Report intakes of foods which are sources of particular carotenoids (Le Marchand et al, 1989) (9).

6. Use the provisional data on carotenoids, and impute values for similar foods (Beecher et al, 1990) (Unpub. data).

Unfortunately, editors do not always send manuscripts for review to researchers who are knowledgeable about food composition data and sources. Perhaps we should be communicating with them! We could suggest that authors presenting dietary intake data follow a prescribed set of rules, similar to the following suggested practices:

1. Identify all primary sources in the text and references.

2. Indicate the source of nutrient data of the selected software package.

3. Describe briefly the methods used for the analysis of particular nutrients.

4. Summarize the procedures used to impute any nutrient values.

I found several papers that satisfied these criteria. For example, Murphy et al. (11) stated that intakes "were calculated using the USDA Survey Nutrient Data Base, which was developed specifically for the NFCS 1977-78 data. The nutrient data were obtained from Agriculture Handbook No. 8, its first three revisions, and additional data. When analytic data were unavailable, nutrient values were derived from similar foods or calculated from recipes; there were no missing values in the nutrient data base." References were provided for all of these sources. This description indicates that others who might wish to repeat this study in another population could essentially repeat the same methods to achieve comparability.

Le Marchand et al. (9) also noted the authors and source of all major references and indicated precisely how the carotenoid content of foods was calculated. Similarly, Byers et al. (8) identified all sources used for analysis of dietary data in their paper.

It is not difficult to identify the source of nutrient values in a database. To illustrate, we are conducting dietary studies among populations living in several diverse geographic areas, and our food composition database needs to include items consumed by multiple ethnic groups in Hawaii, the U.S. Mainland, several Pacific Islands, and Asian countries. Consequently, various published tables and research papers, as well as particular laboratory analyses, are utilized for the database. The sources are identified with a particular code for each nutrient in all food items. Similarly, each diet history questionnaire utilized in a study requires modification of our database so that the values of each food group are representative of the usual diets of the study population. Such alterations should be described in the methodology of the published paper.

As the number of diet-disease studies, databases, and software packages increase, there needs to be better documentation of the source of the food composition data, along with adequate descriptions of the dietary methodology. This is essential for proper interpretation of the study findings and comparability with other published studies.

## REFERENCES


1. Roberts-Gray C, Sparkman AF, Simmons LF, et al. Evaluation of Texas' child-care oriented nutrition education and training program. J Nutr Educ 21:16-25, 1989.

2. Hermann JR, Kopel BH, McCrory ML, et al. Effect of a cooperative extension nutrition and exercise program for older adults on nutrition knowledge, dietary intake, anthropometric measurements, and serum lipids. J Nutr Educ 22:271-274, 1990.

3. Smith-Porter E and Cook RA. Comparison of nutrient intake levels in older adults as estimated by self-administered versus interview-assisted food frequency questionnaires. J Nutr Educ 22:298-302, 1990.

4. Skinner JD. Changes in students' dietary behavior during a college nutrition course. J Nutr Educ 23:72-75, 1991.

5. Paganini-Hill A, Chao A, Ross RK, et al. Vitamin A, beta-carotene, and the risk of cancer: a prospective study. JNCI 79:443-448, 1987.

6. Ziegler RG. A review of epidemiologic evidence that carotenoids reduce the risk of cancer. J Nutr 119:116-122, 1989.

7. Byers TE, Graham S, Haughey BP, et al. Diet and lung cancer risk: findings from the Western New York diet study. Am J Epidemiol 125:351-363, 1987.

8. Hinds MW, Kolonel LN, Hankin JH, et al. Dietary vitamin A, carotene, vitamin C and risk of lung cancer in Hawaii. Am J Epidemiol 119:227-237, 1984.

9. Le Marchand L, Yoshizawa CN, Kolonel LN, et al. Vegetable consumption and lung cancer risk: a population-based case-control study in Hawaii. JNCI 81:1158-1164, 1989.

10. Paul AA and Southgate DAT. McCance and Widdowson's The Composition of Foods, 4th ed. New York: Elsevier/North-Holland, 1978.

11. Murphy SP, Davis MA, Neuhaus JM, et al. Factors influencing the dietary adequacy and energy intake of older Americans. J Nutr Educ 22:284-291, 1990.

# Publishing in the *Journal of Food Compsition and Analysis*

**Kent K. Stewart**
**VPI & SU**
**Blacksburg, Virginia**

The *Journal of Food Composition and Analysis* is an international research journal which is devoted to all scientific aspects of the chemical composition of human foods. The journal emphasizes new methods of analysis; data on composition of foods; studies on the manipulation, storage, distribution, and use of food composition data; and studies on the statistics and distribution of such data and data systems. The journal is looking to build on its strong base in nutrient composition and to place increasing emphasis on other food components such as anti-carcinogens, natural toxicants, flavors, colors, functional additives, pesticides, agricultural chemicals, heavy metals, general environmental contaminants, and chemical and biochemical toxicants of microbiological origin.

My review of the reviewer criticisms of manuscripts over the past few years has lead to the following general list of requirements for manuscript acceptance for the *Journal of Food Composition and Analysis*. Acceptable manuscripts need to have the following attributes.

* **Manuscript Topic Is within Journal's Scope.**

The manuscripts must be original scientific papers or critical scientific reviews about the chemical composition of human food which will be of interest to the readers.

* **Adequate Sample Description**

The foods assayed must be properly described. Key components of the sample description are the number of samples (a composite is one sample), the number of units in a composite, and the number of aliquots analyzed. The author's should provide information on the origin of samples, the places of growth and processing, the point of purchase, and the year of purchase. The cultivar of plant and animal food materials should be provided where possible. Information on the commercial and "home" processing should be provided when they are pertinent to the study. If the food was stored after purchase or processing those storage conditions should be described. In some cases, the composition of a given food is processing and recipe dependent and in such cases it may well be necessary to provide the brand name of the individual food to adequately describe the food.

* **Adequate Description of the Analytical Methodology**

The analytical methodology used for the assays must be referenced. Since the journal has an international readership and some of the world literature is not readily available to some readers, the analytical methodology should be briefly described in the manuscript. Given the reoccurring problems of analytical errors, analytical quality control procedures must be used with the food assays and their use must be described in the manuscript.

* **Comparisons with Previously Published Studies**

The journal readers will almost inevitably want to compare the results of the current study with those of previous studies on the same or similar foods, or with other methods reported to measure the same or similar components. It is the responsibility of the authors of the manuscripts to provide an adequate comparison of their results with those of the pertinent published literature.

It is my experience that careful attention by authors to the issues surrounding these general areas usually results in improved reviewer responses to manuscripts. Failure to properly deal with these issues usually leads to manuscript rejection.

The editorial staff looks forward to receiving your original manuscripts on the chemical composition of human foods for publication in the *Journal of Food Composition and Analysis*. It is hoped that this brief discussion will be helpful in getting your manuscript accepted and published in the Journal.

# Discussion Following Panelist's Presentations

**Grace J. Petot, Moderator**
**Case Western Reserve University**
**Cleveland, Ohio**

The first question which arose during the discussion was "Does anybody care?" Concensus was that, we do care and should be working at finding some solutions or providing guidelines.

The following are comments or responses from panel members and rom conference attendees:

- Specific documentation would allow comparability of studies and/or meta analyses, especially for large epidemiological studies.
- Documentation should be sufficient to allow reproducibility of the work.
- More detail is needed in documenting food frequency questionnaires.
- If a version number, or date, of a system is used, and missing values are added by the user - how should the additions be documented?
- All sources of data should be provided.
- Journal editors should consider selection of reviewers who are knowledgeable about these issues.
- Our levels of concern may need to be matched with the kinds, types, methods, quality of the studies.
- Consider what is reasonable - for editors, users, readers, types of studies, publication space, etc.
- How do we communicate our concerns and recommendations to the broad community?

As an outcome of these presentations and discussion, a task forcewas formed to address the issues. The task force is comprised of the panel members (J. Hankin, E. Monsen, G. Petot, K. Stewart) with Suzanne Murphy as chair.

PART FOURTEEN

# Appendicies

# List of Posters and Poster Abstracts

## SIXTEENTH NATIONAL NUTRIENT DATABANK CONFERENCE POSTER SESSION ABSTRACTS

**An Automated Process for Grouping Food Consumption Data**
D.A. Cook and J.E. Friday

**The Effects of a Computerized Nutrition Program for Senior Citizens**
Darwin Dennison, Kathryn F. Dennison and Joan Y. Ward

**From a Nutrient Data Base to a Food Guide**
Lois Fulton, Anne Shaw and Catherine Tarone

**Computer-Assisted Diet Analysis in an Introductory Nutrition Course**
Francine Genta, N. Jean Downes and Kathryn Sucher

**The Critical Evaluation of Values of Carotenoids for Foods**
J. Holden, A.R. Mangels, G. Beecher, M. Forman and E. Lanza

**The Estimation of Censored Data Points in Small Data Sets**
Joanne M. Holden and Larry Douglas

**Nutrient Density of Minerals and Trace Elements in Diets of Men and Women**
B.S. Hoverson, J.R. Hunt, L.K. Johnson and P.E. Johnson

**International Interface Standard for Food Databases**
J.A.T. Pennington and T.C. Hendricks

**Missing Food Coding Details in a Long-Term, Multicentered Clinical Trial: The First Two Years' Experience**
M.E. Yamamoto, F.A. Averbach, A.W. Caggiula, B.G. Gillis, F.L. Jones, R.M. Meehan, J.A. Naujelis and the MDRD Study

## An Automated Process for Grouping Food Consumption Data

HNIS has under development a modular Food Grouping System that will automate the processes required to link foods reported in USDA's food consumption surveys to recipes for those foods, separate each food into its ingredients, and regroup its ingredients by selected characteristics for analysis. The system can be used to estimate consumption of specific foods, ingredients, or agricultural commodities.

The Food Grouping System will encompass the process of systematically defining, extracting (selecting, converting, linking, combining) verifying, and tracking food and nutrient data at different levels of detail as required for specific research objectives. It will have three distinct functions: aggregation, disaggregation, and normalization. Aggregation draws together into groups; disaggregation separates food groups into components or composite foods and mixtures into their ingredients; and normalization converts data to comparable units. Food grouping is an interdependent and far-reaching process which permeates activities from the defining of food codes to the organization of data for research and reporting.

The HNIS Automated Food Grouping capability will (a) translate intake data on food mixtures into data on ingredients for analyses, tracking, and reporting, (b) facilitate the grouping of food and nutrient intake data from the Nationwide Food Consumption Surveys and (c) accommodate relevant data on survey respondents, such as age, sex and income.

## The Effects of A Computerized Nutrition Program for Senior Citizens[1]

The purpose of the study was to determine the effects of a microcomputerized nutrition program upon the nutrient intake of senior citizens living in residential centers and whether or not hands-on microcomputer interaction improved program satisfaction and compliance motivation. Thirty-one senior citizens between ages 61 and 85 were divided into two experimental groups and a control group. The experimental groups included seniors who attended a nutrition program with hands-on microcomputer interaction (Group 1), and seniors who attended a nutrition program without hands-on microcomputer interaction (Group 2). The control group consisted of senior who did not attend the program. The senior in all groups completed three-day food records at pretest and one month following the program for analysis of their nutrient intake. The seniors in the two experimental groups completed a program satisfaction questionnaire and focus interviews at follow-up. The expected outcomes were that the seniors in Group 1 and Group 2 would have better nutrient intakes than those who were in the control group; and the seniors in Group 1 would be more satisfied with the program and motivated to eat better than the seniors in Group 2. Analysis of variance and t-tests were used to determine differences in nutrient intake and satisfaction/motivation scores. The findings indicated that the nutrition program had a significant effect on the intake of saturated fat at follow-up for both experimental groups. The seniors in Group 1 were more satisfied with the program that the seniors in Group 2. Nutrient analysis methodology, if integrated properly, personalizes nutrition education programs, improves participant satisfaction, and enhances motivation to comply.

[1] Funded, in part, by the National Institute on Aging, R 43, 44AG 06269-02A2 Computerized Nutrition Program for Senior Citizens, May 1990

## From a Nutrient Data Base to a Food Guide

USDA's research-based food guidance system is based on what foods Americans eat, what nutrients are in these foods, and how to make the best food choices. The food guidance system provides a basis for a total diet that contains the nutrients needed but not too many calories or too much fat, saturated fat, cholesterol, sugar, sodium, or alcohol.

The food guidance system is supported by a framework or core of major groups of nutrient-bearing foods with suggested ranges of servings from each group. The objectives for the system specified a range of food energy and nutrient levels, moderate levels of fat, cholesterol, sweeteners, and sodium, and adequate levels of complex carbohydrate. Patterns for diets at specific calorie level using this framework of suggested servings from the major food groups and subgroups were defined to meet a range of energy levels. Ability of the patterns to meet nutrient objectives was evaluated using food composites for each group and subgroup. These composites were based on typical selections of foods in each group as reported by individuals in the 1977-78 Nationwide Food Consumption Survey. For example, a food chosen more frequently was represented as a greater proportion of the composite. For the purpose of this evaluation, foods in the composites were generally considered to be in their leanest forms, such as nonfat milk, meats trimmed of all fat, poultry without skin, and fruits and vegetables without added fat(s) or sweetener(s).

Nutrient content of each composite was determined using the Nutrient Data Base for Individual Food Intake Surveys, Release 4.0. Nutrient content of the foods in a pattern was calculated by multiplying the number of servings of each composite and then obtaining a total amount of each nutrient. Nutrient levels in the patterns were compared to Recommended Dietary Allowances for the sex-age groups whose energy needs fit the calorie level of the pattern to assess nutrient adequacy of the patterns.

## Computer-Assisted Diet Analysis in an Introductory Nutrition Course

We compared computer-assisted and manual calculation methods of diet analysis in an introductory nutrition course. A blind crossover design grouped students by order of exposure to method (COMPFIRST OR HANDFIRST). Nutrition knowledge was assessed after completion of the first diet analysis assignment. In order to assess attitudes, students were permitted to select their preferred method of analysis to evaluate their own diets, and were grouped (CHOSECOMP OR CHOSEHAND) according to the method chosen. Comparisons of the COMPFIRST (n=43) and HANDFIRST (n=51) groups by Student's t-test revealed no significant differences in nutrition knowledge as measured by exam scores following the first assignment. Students in both the CHOSECOMP (n=41) and CHOSEHAND (n=48) groups were significantly more likely ($p \leq .0001$) in the CHOSECOMP group. Both teacher and students reported the computer method saved time. Student selection of the manual calculation (CHOSEHAND) method was related to flexibility in scheduling work place and time and not to any educational differences between methods.

## The Critical Evaluation of Values of Carotenoids for Foods

The inverse association between the dietary intakes of fruits and vegetables and the risk of certain cancers rests, in part, on the biological model of carotenoids as anti-oxidants. However, the test of the association between the dietary intake of specific carotenoids and cancer incidence requires the availability of a carotenoid food composition database. The authors compiled values for alpha-carotene, beta-carotene, beta-cryptoxanthin, lutei+zeathanin, and lycopene in foods from over 100 published and selected unpublished sources. An expert system was developed to rate each value on the basis of 1) analytical method, 2) quality control procedures, 3) sampling plan, 4) sample handling, and 5) number of samples analyzed. Following the evaluation of data, acceptable values for each carotenoid and food were combined to generate the CAREX database which contains more than 250 foods. The database includes the food description, the mean value for each carotenoid, the number of acceptable values, acceptable references used, and a confidence code. The confidence code is derived from the ratings and is an indicator of the reliability of a specific carotenoid value for a food. This database is a useful tool for the evaluation of the intake of specific carotenoids in epidemiological studies of cancer incidence.

## The Estimation of Censored Data Points in Small Data Sets

The accurate estimation of low levels of components in foods is difficult duet to the censoring of a distribution at the lower end or "left-censoring" as a result of the quantitation limitation of analytical methods. Several techniques using constant values as well as conditional maximum likelihood estimators were evaluated in samples of 5, 10, and 20 observations taken from simulated populations. Samples from normal, log-normal, and square-root distributions were considered. In all cases, percent bias and the coefficient of variation for standard errors of the data set means were smaller when conditional maximum likelihood estimators were used.

## Nutrient Density of Minerals and Trace Elements in Diets of Men and Women

Healthy, free-living men and women (n=107), 20 to 59 years of age, who were not taking medication or dietary supplements were studied to determine whether the nutrient density of minerals and trace elements of foods consumed varied by age and sex. Nutrient intakes were estimated from 3-day food records. Mean (±SD) caloric intake was 1731 ± 358 for women and 2420 ± 517 for men. Women consumed 25.8 ± 7.9 kcal per kg body weight which was significantly less than men who consumed 30.0 ± 8.4 kcal per kg body weight ($p<0.01$). The body mass index did not differ between groups (26 ± 6 vs 27 ± 5 kg/m$^2$, respectively). Mineral and trace element intakes follow:

| | Mg/Day Women | Men | Mg/1000 Kcal Women | Men | Mg/kg Body Wt Women | Men |
|---|---|---|---|---|---|---|
| Ca | 787±272 | 945±425* | 464±147 | 389±145* | 11.7±5.0 | 11.7±5.6 |
| Cu | 1.0±0.2 | 1.3±0.4** | 0.6±0.2 | 0.5±0.1** | 0.02±0.01 | 0.02±0.01 |
| Fe | 12.8±5.7 | 14.9±5.2* | 7.6±3.2 | 6.3±2.0** | 0.2±0.1 | 0.2±0.1 |
| Mg | 263±64 | 316±88** | 156±39 | 133±36** | 3.9±1.1 | 3.9±1.3 |
| P | 1190±310 | 1477±424** | 696±139 | 614±138** | 17.6±5.7 | 18.2±5.4 |
| Zn | 9.9±3.0 | 12.4±3.2** | 5.8±1.6 | 5.2±1.2* | 0.2±0.1 | 0.2±0.1 |

*p>0.05 **p<0.01

Age did not affect intakes of minerals and trace elements for men or women. Dietary intakes of all nutrients were greater for men than women, however, nutrient densities (mg/1000kcal) were greater for women. Nutrient intakes/kg body weight were similar for men and women. Thus, the women reported diets that were more concentrated in minerals and trace elements, resulting in nutrient intakes/kg body weight remarkably similar to those of the men.

## International Interface Standard for Food Databases

A database interface protocol incorporating LANGUAL as the coding scheme and retrieval methodology will be developed to allow easy access to and convenient international exchange of food consumption data. LANGUAL is an indexing language which allows for computerized retrieval of food names (and the data associated with these foods) from databases relative to 13 characteristics that affect the safety and/ or nutritional quality of foods. The past 19 years of work on this system have included the development of a thesaurus (with terms, definitions, and scope notes), a hierarchical tree structure of terms, and a manual that explains the factors, factor terms, and factoring rules to indexers and searchers.

The first step in this project will be to develop, refine, and implement a schema for an international interface standard for food databases. The second step will be to develop a repertoire of personal computer programs for retrieval of food data using LANGUAL and the standard interface. The resulting system will be usable on a range of personal computers. The results of the efforts of this project will be disseminated to the widest possible international audience through demonstration projects for international symposia and pilot tests for actual scientific application.

## Missing Food Coding Details in a Long-Term, Multicentered Clinical Trial: The First Two Years' Experience

The Modification of Diet in Renal Disease (MDRD) Study, a multicenter, five-year clinical trial, is sponsored by the National Institutes of Health and the Health Care Financing Administration and is designed to determine whether controlled dietary protein and phosphorus intake and/or control of blood pressure will alter the progression of chronic renal disease. Completeness of dietary data is a complex quality control challenge in such long duration, collaborative studies. Since dietary assessments depend on the completeness of intake information, the identification and alleviation of missing diet detail problems is key.

To examine patterns of problems in dietary data completeness, we reviewed the first two years of NCC clarification requests ("queries") for missing food record details. Clarification requests increased from 14.6% of 6,287 recorded days in 1989 to 25.4% of 8,687 days in 1990. Some improvements in quantity descriptions were noted. Overall, however, similar levels and types of problems for specific foods continued over the two years. For example, missing details on bakery goods, poultry and supplements. The rise in clarification requests was predictable, in some items; in the volume of food records for fixed staff levels and in proportion which was follow-up data. In long-term studies, reviews such as this provide valuable insights into needed support for patients and dietitians' work with research diet data.

# List of Exhibitors

| Company | Conference Contacts |
|---|---|
| **The CBORD Group**<br>P.O.Box 700<br>Ithaca, NY 14851<br>607-257-2410 | Ann Anderson<br>Nancy Vergara |
| **DINE Systems**<br>586 N. French Road<br>Suite 2<br>West Amhurst, NY 14228<br>716-688-2492 | Joan Y. Ward<br>Darwin Dennison |
| **ESHA Research**<br>P.O. Box 13028<br>Salem, OR 97309<br>503-585-6242 | Elizabeth Hands<br>Patricia Bishop |
| **formulas/NOW!**<br>2755 Chelsea Drive<br>Suite A<br>Oakland, CA 94611-2505<br>415-482-2705 | Jerry Salzman |
| **N-Squared Computing**<br>3040 Commercial St. SE<br>Suite 240<br>Salem, OR 97302<br>503-364-9118 | Laurie North |
| **Nutrition Coordinating Center**<br>University of Minnesota<br>2221 University Ave. SE<br>Suite 310<br>Minneapolis, MN 55414<br>612-627-4862 | Mary Stevens |
| **Pages Software**<br>2070 30th Avenue<br>San Francisco, CA 94116<br>415-369-0377 | Pat Booth<br>Glenn Lew |
| **Technical Assessment Systems**<br>1000 Patomac Street, N.W.<br>Suite 102<br>Washington, DC 20007<br>202-337-2625 | Judi Douglass |
| **University of California Berkeley**<br>Nutritional Sciences Dept<br>9B Morgan Hall, UC<br>Berkeley, CA 94720<br>415-643-6423 | Karl Grose |
| **University of Texas**<br>Health Science Center<br>School of Public Health<br>P.O. Box 20186, Room W606<br>Houston, TX 77225<br>713-792-4660 | Dierdre Douglass |
| **USDA, HNIS**<br>Nutrition Monitoring Division<br>Federal Building, Room 304<br>Hyattsville, MD 20782<br>301-436-8507 | David Haytowitz |
| **Wellsource**<br>15431 SE 82nd Drive, Suite D<br>P.O. Box 569<br>Clackamas, OR 97015<br>503-656-7446 | John Gobble |
| **Whole Grain Software**<br>1427 Bancroft Way<br>Berkeley, CA 94702<br>415-848-5903/5972 | David Stone |