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.

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

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.

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.