8th National Nutrient Databank Conference

July 25-27, 1983
Minneapolis, Minnesota

Data Base Building and Use

- Microcomputer-based Interactive Methods to Collect and Code Dietary Data
  I. Marilyn Buzzard, Nutrition Coding Center, University of Minnesota
- The Highs and Lows of Building a Data Base
  Judith E. Brown, Public Health Nutrition, University of Minnesota
- Nutrient Data Bases in Computer Programs for Undergraduate Food Sci\ence and Nutrition Curricula
  Elaine H. Asp, Department of Food Science and Nutrition, University of Minnesota

Keynote Speaker:

Future Computers: New Perspectives for Nutrient Databanks
Earl C. Joseph, Futurist, Anticipatory Sciences, Inc.

User Workshops

- Computer Systems
  - James J. Doyle, Honeywell, Datanetwork
  - Loretta W. Hoover, University of Missouri-Columbia
- Patient Care - Dietary Evaluation
  - Ann T. Sandvick, Children's Hospital, Medical Center
  - Joyce B. Wenz, University of Minnesota
- Food Service--Menu and Recipe Planning
  - Linda S. Hicks, St. Luke's Episcopal Hospital/Texas Children's Hospital
  - Fay Choban, University of Minnesota
- Data Base Development
  - Lillian White, Dietary Analysis and Prescription Services, Inc.
  - Jean Pennington, Food and Drug Administration
- Education
  - Elaine H. Asp, University of Minnesota
  - Joanne L. SlaVin, University of Minnesota
- Industrial Uses
  - Rose Tobelmann, General Mills, Inc.
  - Barbara Bentson, The Pillsbury Company

Common Conventions

- Commonalities and Differences: Are We All Speaking the Same language?
  Grace J. Petot, Case western Reserve University
- Report of International Meeting
  R. Guarth Hansen, Utah State University
- Standards Development: The ASTM Way
  John J. Rothrock and Kenneth C. Pearson, American Society for Testing and Materials

FROM THE EDITOR

The Eighth National Nutrient Data Bank Conference was held from July 25-27, 1983 in Minneapolis, Minnesota. The Proceedings for the conference were developed from the speakers presentations at the conference. Because of the cooperation of all the speakers, I am proud to say we have an excellent record of the conference for the conference attendees and other persons interested in nutrient data bases.

The editor is indebted to the conference Arrangements Committee and Dr. Mary Ann Smith of the Office of Special Programs at the University of Minnesota for their help and guidance in making the conference a success. In addition, I would like to thank Jean Shields for her organizational assistance in arranging meetings and typing letters for the conference. Finally, I would like to express my sincere gratitude to Karen Schmalz who, with the wizardry of Wang, was able to compile these proceedings from a pile of disorganized paper. These secretaries deserve credit for coping with my unlimited requests related to the conference.

Rose Tobelmann
Editor, Conference Proceedings

1983 NUTRIENT DATA BANK CONFERENCE

This conference was made possible in part by generous contributions from these companies and institutions:

United States Department of Agriculture

National Technical Information Service

Best Foods

Campbell Institute for Research & Technology

The Doyle Pharmaceutical Company

Frito-Lay, Incorporated

General Foods Corporation

General Mills, Incorporated

Hershey Foods Corporation

Hoffman-LaRoche, Incorporated

Mead Johnson, Nutritional Division

The Pillsbury Company

Ross Laboratories

EIGHTH NATIONAL NUTRIENT DATABANK CONFERENCE

MINNEAPOLIS, MINNESOTA

JULY 25-27, 1983

COMMITTEES

Steering Committee

Al Riley, Chairman
Jim Doyle
Loretta Hoover
Betty Perloff
Ken Samonds
Jack Smith
Rose Tobelmann

Program Committee

Jack Smith, Chairman
Linda McAllister
Grace Petot
Ann Sorenson

Arrangements Committee

Rose Tobelmann, Chairman
Elaine Asp
Howard Bauman
Joyce Wenz

Communications Committee

Betty Perloff, Chairman
Suzanne Pelican

Exhibits Committee

Jim Doyle, Chairman
Joann Slavin

Databases Committee

Loretta Hoover, Chairman
Tony Fisher
Donna Hay

Pre-Conference Committee

Ken Samonds, Chairman
Linda Campitelli
Lilly Gardner
Stan Gray

NUTRIENT DATA BASES: AVAILABILITY, OPTIONS AND RELIABILITY

Betty Perloff
Consumer Nutrition Division
Human Nutrition Information Service, USDA
Hyattsville, Maryland 20782

The key parts to any Nutrient Data Base System are the computer, the programs (software), and the nutrient data base. The nutrient data base is the collection of nutrient values that is accessed by the programs as nutrient values are needed for calculations.

AVAILABILITY AND OPTIONS

Nutrient data bases currently available from the U.S. Department of Agriculture (USDA) either correspond to a USDA publication, or were developed to be used with the USDA Nationwide Food Consumption Survey. Agriculture Handbook No. 8 is the Department's standard reference on the composition of foods. A complete revision of the handbook was published in 1963 (Watt and Merrill, 1963). It is now undergoing another revision which is being completed and released in separate sections according to food groups. Nine sections have been released to date (U.S. Department of Agriculture, 1976-1982), and an additional 14 sections will be published over the next few years. The nutrients covered in the current revision and the earlier edition are shown in Table 1.

Table 2 lists computerized data sets of nutrient values (nutrient data bases) prepared at the Consumer Nutrition Division of USDA's Human Nutrition Information Service. The Nutrient Data Base for Standard Reference corresponds in both content and format to the revised sections of Handbook 8 and contains the largest number of nutrient values. As new sections of the handbook are completed, the data set is updated and a new version is released. Each version is identified by a release number and date. Until the revision is complete, releases of this data base will be supplemented with data from the earlier edition of the handbook. Of course, data from the earlier handbook are limited to the nutrients that were reported in that edition.

Data Sets 456-3 and 456 correspond to the 1963 edition of the handbook. They contain identical nutrient data, but in 456-3, data are expressed on the basis of 100-gram edible portions, and in 456, on the basis of household measures. Data Set 72-1 is an abbreviated data base corresponding to Home and Garden Bulletin No. 72, which is a food composition publication intended for consumers' use.

Imputed values have been included in Data Sets 456-3, 456, and 72-1 where blank spaces appeared in the publications. Imputed values have been included in the Standard Reference Data Base for most of the blank spaces in the revised section of the handbook. Exceptions are blank spaces for: spices and herbs; nutrients in food groups where analytical data were unavailable for all foods within the group (e.g., copper in dairy products); and fatty acids, cholesterol, phytosterols, and amino acids for baby foods and foods from plant products. Values for enriched flour and bread and other products made with enriched flour have been updated to reflect the current standards of identity.

The last two data bases listed on Table 2 were developed for the analysis of data from the USDA Nationwide Food Consumption Survey 1977-78. The Data Base for Individual Intake Surveys contains data on the 100 gram basis for those food items reported as being consumed by respondents in the individual phase of the survey. The Data Base for Household Food Use Surveys contains data for 1 pound as purchased for food items reported in the household phase of the survey. Cooking losses were deducted for vitamins in this last data base. Values were imput for both survey data bases where data were unavailable.

These six data bases are available on magnetic tape from the National Technical Information Service (NTIS) at the current cost of $140 per tape. Information about tape formats and instructions for ordering the tapes may be obtained by writing the Consumer Nutrition Division[1]. The USDA data bases do not include computer programs.

The Nutrient Data Base for Standard Reference is used in the Consumer Nutrition Division to create other, more specialized, data bases. Figure 1 shows the interrelationship between the Standard Reference Data Base and the Individual Survey Data Base. The large box represents our Nutrient Data Bank. The three circles within this box represent the three levels of data with the Bank: Level 1 (sometimes called Data Base 1) = individual values; Level 2 (Data Base 2) = preliminary summary; Level 3 (Data Base 3) = final summary.

The food items for which data appear on the Survey Data Base are foods that respondents consumed during the survey reporting period. The items are frequently composites, or recipes, of two or more foods from the Standard Reference file. For food items for which data summarization is incomplete, information may be drawn from the data bank at Level 2. If there are no data available for an item, values are imput or estimated. The food identification numbers on the survey data base are different from those on the Standard Reference Data Base because they were designed to allow for specific data processing needs for the survey.

There are also several commercial data bases available and a directory of these data bases is being prepared by a volunteer committee chaired by Dr. Loretta Hoover. The directory will contain information on all available nutrient data bases, other than USDA data bases, that the committee could locate and for which they could obtain a completed survey questionnaire. Dr. Hoover is making arrangements to have a copy of the directory mailed to each conference registrant, and additional copies may be obtained through the University of Missouri.

The different commercial sources listed in the directory offer different products and services--from data bases that are basically copies of the USDA data, to data bases with greatly expanded coverage. Most of the commercial data bases are accompanied by computer programs that can be used with their respective data bases. Some of these products may be purchased, some may be accessed through a remote terminal. Some of the sources will perform the data entry and computer processing for users who do not have access to a computer or a terminal. The directory summarizes the services that each source offers, and provides the name of a contact person and phone number for obtaining additional information.

---

[1]Consumer Nutrition Division, Human Nutrition Information Service, U.S. Department of Agriculture, Room 304, 6505 Belcrest Road, Hyattsville, MD 20782.

OPTIONS

Summarizing briefly the options that are available for using a computerized nutrient data base, a user may:

- Purchase a data base and prepare the computer programs (the data base may be from USDA, or a commercial source);
- Purchase both the data base and programs;
- Contract to use a system through a remote terminal; or
- Contract with a nutrient data base system for nutrient data calculations.

The choice depends on which option will suit specific needs. Regardless of the option chosen, accuracy of the data and the programs must be of primary concern to the user. Studies comparing nutrient data base systems have shown that different systems may give different results (Danford, 1981; Hoover, 1981). Causes of the differences may be the nutrient data, the computer programs, or coding differences.

There are several reasons why nutrient values may differ from one system to another, even though literature from most systems cite USDA as their data source. One reason is that data from USDA are being updated, and some systems may be using information that has been superseded. Also, systems may supplement their data bases with information from other sources, for food items or nutrients not covered by USDA. If two systems use different sources for supplementary data, the values may differ. Other systems may choose not to supplement their data bases and may instead perform calculations with the data that are present, flagging the calculation results when data for an item are missing.

RELIABILITY

The nutrient content of any food varies naturally from sample to sample, and conditions under which foods are grown or stored, as well as any processing or preparation treatments, may increase the variability. The only way to know the exact nutrient content of a particular food sample is to perform analytical measurements on that sample in a laboratory. Such measurements must be made for metabolic studies or other projects for which precise knowledge is essential. Agriculture Handbook No. 8 provides data on the nutrient content of foods that can be used when direct measurements are not practical.

The process of deriving nutrient values for Agriculture Handbook No. 8, including the current revision, differs as circumstances prescribe, but the goal always is to present values that are representative of products on a year-round, nationwide basis. For that reason, the data published in the handbook are usually referred to as "representative values". Users should not mistakenly interpret this term to infer that the values are not derived from analytical data. Except for fortified nutrients in some foods, all of the values in the handbook are based on analytical data. Most of them are arithmetic means of several individual values.

Values that are not themselves averages of actual analytical data are usually calculated in some particular way from the analytical values. Protein, for example, is not analyzed directly but is calculated from the average nitrogen value for each food. Calculations for some food mixtures are made using analytical values for the ingredients in the recipe, adjusting the total values for changes in moisture or fat content, and for losses in vitamins and minerals, that occur during cooking of the food.

Individual analytical values are collected by USDA from a variety of sources, primarily from the food industry, scientific literature, and government laboratories. Individual values for each food item are stored in USDA's computerized Nutrient Data Bank, organized into subgroups according to the unique characteristics of the individual foods sampled, analyzed or studied to determine the factors that affect each item's nutrient composition, and summarized to obtain the mean values reported in Handbook 8.

A critical factor in the reliability of nutrient data is the analytical methods that are used for measuring the nutrient levels in foods. Dr. Gary Beecher, Chief of USDA's Nutrient Composition Laboratory, recently prepared a chart showing the current status of nutrient methodology (reproduced following Figure 1). According to this chart, most nutrients in Handbook 8 have acceptable methods. Note, however, that some of the nutrients designated as having methods that give conflicting results do appear in the handbook. Ten components are designated as lacking suitable methods; none of these components appear in Handbook 8.

APPRAISING NUTRIENT DATA BASES

Most commercial data bases use one of the USDA data bases as a core, supplementing it with nutrient values from other sources to provide coverage beyond the coverage available from USDA. To evaluate a commercial nutrient data base, first obtain information to answer the following series of questions:

1. Data Base Core

   Which USDA data base is used as the core?

   What was the date of its release?

   Have any revised sections of Agriculture Handbook No. 8 been published since the release date, and if new sections have been published, have the revised values been incorporated into the data base?

   If the release date precedes 1977, have values for B vitamins in enriched bread and flour, and foods using enriched flour, been updated to reflect the current standards of identity for those products? (The standards were revised in 1975. USDA data bases released in 1977 and thereafter were changed accordingly.)

   If the release date precedes 1982, have values for iron in the enriched flour and bakery products been updated? (New standards for iron became mandatory July 1, 1983. The revised iron levels could be used voluntarily by manufacturers before the effective date, and the 1982 USDA data releases reflected this revision.)

2. Supplementary Data

What, if any, data have been added to the data base?

What are the sources for the supplementary data?

Do personnel selecting the supplementary data have a food science, dietetics, or similar background?

What precautions are taken to ensure reliability and accuracy of the supplemental data?

3. Nutrient Coverage

Are any nutrients included in the data base that are not included in Handbook 8? If additional nutrients are included, how complete is their coverage?

Are any of the nutrients in the data base included in revised sections of Handbook 8 but not in the earlier edition, i.e., B_6, folacin, pantothenic acid, magnesium, zinc? Their coverage will be incomplete until all sections of the handbook revision are published. How complete is their coverage in the data base for those food groups that have not yet been revised--that is, how many of the food items actually have values present for these nutrients?

Are any nutrients which have questionable analytical methods included in the data base? (Biotin, choline, molybdenum, vitamin K, cobalt, heme-iron, nonheme-iron, silicon, tin, or vanadium). If any of these nutrients are included in the data base, how complete is their coverage? (Remember, data for these nutrients are very limited and their analytical methods need further study.)

If dietary fiber is included in the data base, is it identified properly? (Dietary fiber consists primarily of cellulose, hemicellulose, lignin, pectin, and gums. Different methodologies, however, measure different combinations of these components. The neutral detergent method has become increasingly popular in the last few years, but it measures only the insoluble components: cellulose, some of the hemicellulose, and lignin. These data should be identified as insoluble dietary fiber, or as neutral detergent fiber. They do not represent the total dietary fiber in foods containing soluble components.)

4. Management

Is information about the data base core, past updates, and inclusion of supplementary data well-documented and readily available?

Is the data base manager familiar with Agriculture Handbook No. 8, including the status of the current revision?

Once answers are assembled for the above questions, they can be used to judge the status of the data base, the completeness of the nutrient coverage, and whether sufficient effort is being made by the data base management to provide a high quality product.

RESPONSIBILITIES OF A DATA BASE OWNER

Once a data base is selected and purchased, new responsibilities are acquired if the data base is to be used over a period of time. The data must be updated as new releases from USDA are made. If data are to be added from other sources, criteria for data selection should be established. All updates and additions should be completely documented. Whenever the data base is used in a research project, the name, release number, and release date of the USDA core data base, as well as information about updates and added data, should be included in the research report.

Responsibilities also extend to ensuring the accuracy of computer programs that are written for use with the data base. Any program that you purchase, that you prepare yourself, or that is prepared for you by someone else should be checked against a set of hand calculations.

It is also important for a system's users to understand the calculating procedures that are employed by the system. Our office has cooperated with the University of Missouri at Columbia to develop a model[2] (Hoover and Perloff, 1981) that can help a user review a nutrient data base system. The model has three main parts. The first part contains questions for which any user should know the answers relative to any system that he/she is using. Second, there is a series of five computer processing tasks that must be completed using the nutrient analysis system. These include: updating the data base, calculating a simple recipe, reporting baseline data, converting nutrients to various portion sizes, and computing a dietary record. The third part of the model is an interpretation guide to facilitate evaluation of the results from the computing tasks; it includes reference data calculated from USDA sources for comparison with the computer output, as well as several questions which should be answered during the computer output evaluation. The model was prepared as a prototype and may not suit all purposes, but individuals are free to utilize the methodology to prepare a more specialized tool for their own purposes.

CONCLUSION

Many decisions face potential nutrient data base users. The first step toward becoming a responsible user is to recognize that all nutrient values, if not direct analyses of the particular samples under question, are representative values, and some are more reliable than others. The next step is to accept the responsibility for the reliability of the nutrient data base being used, and to take the necessary precautions to ensure that it is current and accurate.

---

[2] "Model for Review of Nutrient Data Base System Capabilities" may be ordered from Department of Human Nutrition, Foods and Food Systems Management, College of Home Economics, 217 Gwynn Hall, University of Missouri-Columbia, Columbia, Missouri 65211. Price $7.50.

LITERATURE CITED

Danford, D.E., 1981. Computer applications to medical nutrition problems. JPEN 5:441-446.

Hoover, L.W,, 1983. Computerized nutrient data bases: I. Comparison of nutrient analysis systems. J. Am. Dietetic. A. 82:501-505.

Hoover, L.W. and B.P. Perloff, 1981. Model for Review of Nutrient Data Base System Capabilities. Columbia: University of Missouri-Columbia Printing Services.

U.S. Department of Agriculture, 1976-1982. Composition of Foods: Raw, Processed, Prepared. USDA Agriculture. Handbook No. 8. Revised Sections: 8-1, Dairy and Egg Products; 8-2, Spices and Herbs; 8-3, Baby Foods; 8-4, Fats and Oils; 8-5, Poultry Products; 8-6, Soups, Sauces, and Gravies; 8-7, Sausages and Luncheon Meats; 8-8, Breakfast Cereals; and 8-9, Fruits and Fruit Juices.

Watt, B.K. and A.L. Merrill, 1963. Composition of Foods...raw, processed, prepared. Rev. USDA Agriculture Handbook No. 8.

Table 1. Nutrients in Agriculture Handbook No. 8

	1963 edition	Current revision
Proximates:	Moisture	Moisture
	Energy	Energy
	Protein	Protein
	Total lipid	Total lipid
	Carbohydrate	Carbohydrate
	Crude fiber	Crude fiber
	Ash	Ash
		Dietary fiber, insoluble *
Minerals:	Calcium	Calcium
	Iron	Iron
	Magnesium *	Magnesium
	Phosphorus	Phosphorus
	Potassium	Potassium
	Sodium	Sodium
		Copper *
		Manganese *
		Zinc
Vitamins:	Ascorbic acid	Ascorbic acid
	Thiamin	Thiamin
	Riboflavin	Riboflavin
	Niacin	Niacin
	Vitamin A	Vitamin A
		Pantothenic acid
		Vitamin B_6
		Folacin
		Vitamin B_{12}
		Tocopherol *
Lipids:	Cholesterol *	Cholesterol
	Oleic acid *	Oleic acid
	Linoleic acid *	Linoleic acid
	Total saturated fatty acids *	Total saturated fatty acids
		Total monounsaturated fatty acids
		Total polyunsaturated fatty acids
		Other fatty acids
		Phytosterols *
		Amino acids

*Limited data provided.

Table 2. USDA computerized nutrient data bases

USDA Nutrient Data Base for Standard Reference, release 3, 1983 *
Data Set 456-3, release 4, 1983 +
Data Set 456, release 4, 1983 +
Data Set 72-1, release 2, 1982
USDA Nutrient Data Base for Individual Intake Surveys, release 1, 1980
USDA Nutrient Data Base for Household Food Use Surveys, release 1, 1980

* Release 3 includes information from revised sections 8-1 through 8-9
\+ Updated values from revised sections 8-1 through 8-9 have replaced original data for those food groups

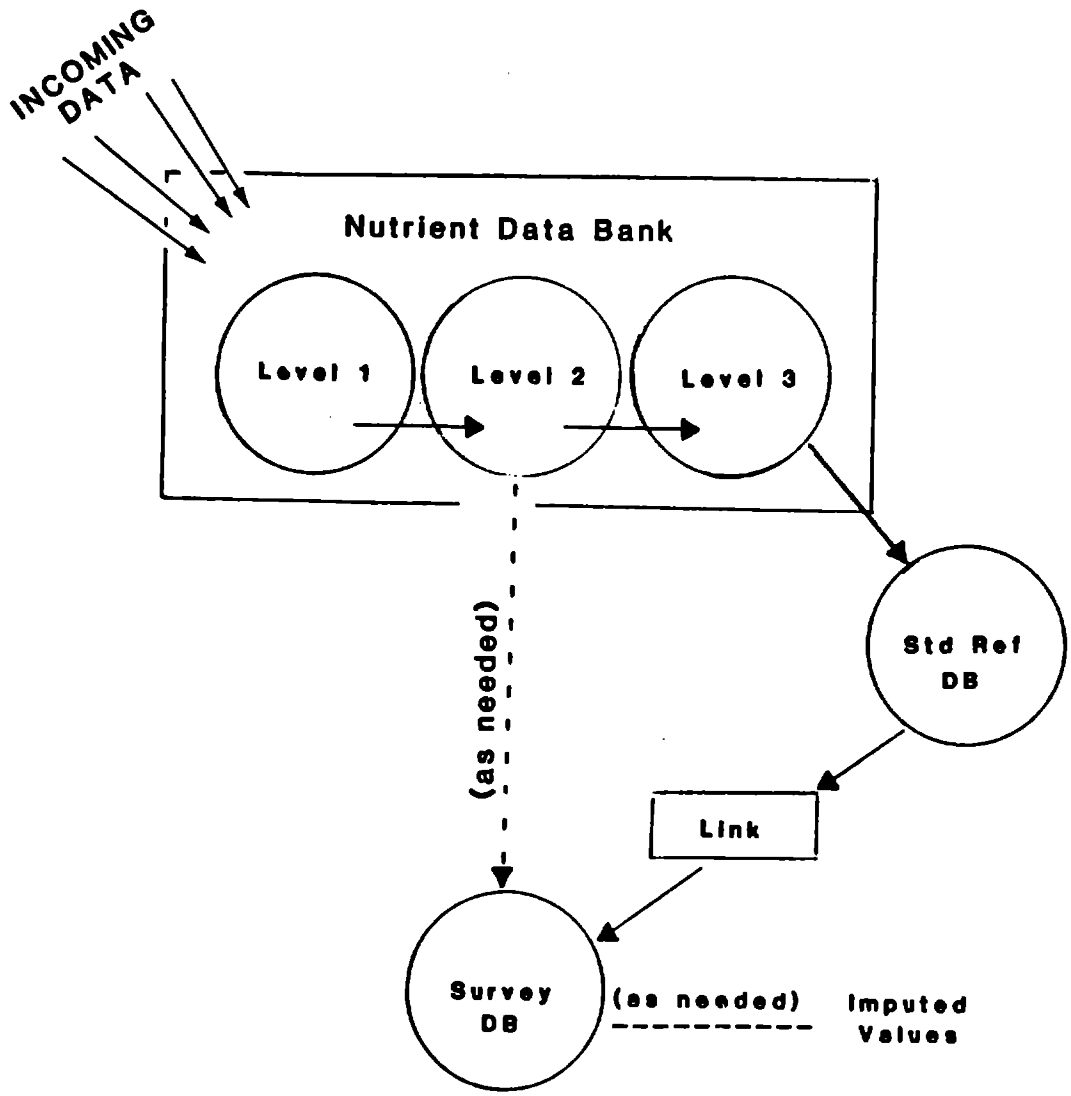

INCOMING DATA
Nutrient Data Bank
Level 1
Level 2
Level 3
(as needed)
Std Ref DB
Link
Survey DB
(as needed)
Imputed Values

Figure 1

STATE OF DEVELOPMENT OF METHODS FOR NUTRIENTS IN FOODS

Nutrient Composition Laboratory
ARS, USDA
Beltsville, MD 20705
July 1983

	State of Methodology[a/]			
Nutrient category	Adequate	Substantial	Conflicting	Lacking
Carbohydrates, fiber and sugars		Individual sugars	Fiber Starch	
Energy			Food energy	
Lipids		Cholesterol Fat (total) Fatty acids (common)	Sterols Trans-fatty acids	
Minerals/Inorganic nutrients	Calcium Copper Magnesium Phosphorus Potassium Sodium Zinc	Iron (total) Selenium	Arsenic Chromium Fluorine Iodine Manganese	Cobalt Heme-iron Molybdenum Nonheme-iron Silicon Tin Vanadium
Proteins and amino acids	Nitrogen (total)	Amino acids (most)	Amino acids (some) Protein (total)	
Vitamins		Niacin Riboflavin Thiamin Vitamin B-6	Vitamin A Carotenes Vitamin B-12 Vitamin C Vitamin D Vitamin E Folacin Pantothenic acid	Biotin Choline Vitamin K

a/ Description of methodology states

Factors	Adequate	Substantial	Conflicting	Lacking
Accuracy	Excellent	Good	Fair	Poor
Speed of analysis	Fast	Moderate	Slow	Slow
Cost per analysis	Modest (<$100)	Modest to high	High	?
Development needs	---	Method modif. Extraction proc. Applications	Method develop. modif. Extraction proc. Applications	Method develop. Extraction proc. Applications

From: Beecher, G.R. and Vanderslice, J.T. Determination of Nutrients in Foods. Chapter in Proceedings of Symposium on Modern Methods of Food Analysis. Editors K.K. Stewart and J.R. Whitaker. Westport: AVI, 1984 (in press).

CODING AND ENTRY OF FOOD INTAKES

Lenore Arab, Ph. D.
Klinisches Institute für Herzinfarktforschung
der Medizinischen Universityätsklinik
Bergheimer Str. 58
D-6900 Heidelberg, FRG

Those more than 250 of you attending this pre-conference either use or wish to use nutrient data bases to convert food intakes or recipes into nutrients. Such usage requires a nutrient composition source, access to a computer and also a way of communicating with the computer. That is, a way of telling her which foods you are interested in. This presentation focuses on this communication problem--how to code and enter foods for analysis with a computerized nutrient data base. There is no single way of doing this; the theory behind this, various approaches to coding and entry, and the considerations and complications which guide in the selection of a method will be presented. Finally, how these tasks are performed in Heidelberg will be discussed.

How did I become involved? I am working in the field of nutritional epidemiology in the FRG and conduct nutrition and health surveys quite like a German HANES--but with different emphasis, and different dietary methodology (1)--we take great pains to collect seven-day dietary records from our subjects--which requires entry and coding of massive amounts of dietary information, and is the most labor intensive part of our clinical examinations.

Coding is always an expensive, labor intensive part of nutriental analysis, and a large potential source of error, but also as well, a critically important part of nutrient analyses--and too often neglected. Coding can be simplified, but not eliminated, by modern technology. It is a process, and not a single step.

Coding is the conversion of the information on items eaten or to be eaten into machine readable form.

Data entering is forcing the machine to read the coded information. Together they achieve what is being attempted in Figure 1; communication between subject and computer about what has been eaten.

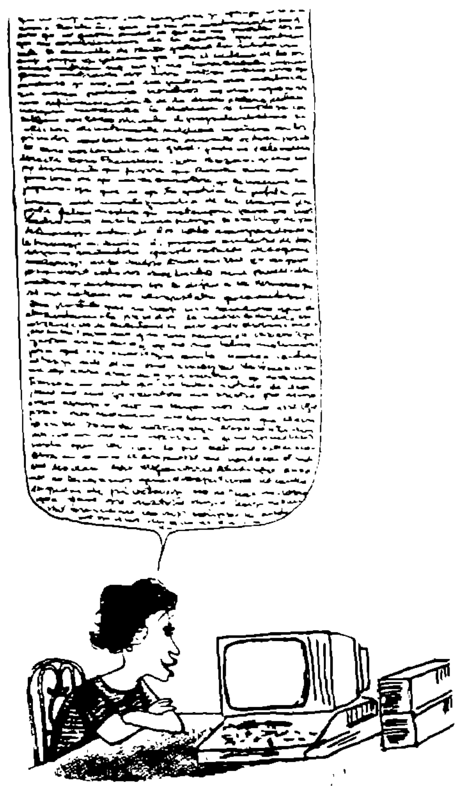


Figure 1.

Let's take a look at the steps involved in the calculation of a 24 hour recall to get a feel for what coding and entry entail. The steps with the orientation on the personnel normally required in this process include someone to interview; that is ask the subject exactly what and in which quantity was consumed yesterday; someone to look up and transcribe the codes for these foods and determine their amounts; someone to punch this information. This information must be fed into the machine by someone (leading to possible reading and typing errors); then the information in the computer should be checked for correctness against the originals; and the inevitable mistakes in such a system corrected, that is rewritten, typed, entered and checked. These steps and the dangers associated with each threatening the accuracy of the data are detailed in Table 1.

TABLE 1

STEPS INVOLVED IN 24 HOUR DIETARY RECALL ASSESSMENT

Steps	Potential Error Sources
1. Question person about what was eaten	Memory failure Influence of interviewer
2. Question about how much was eaten	Memory failure Estimation difficulties Non-edible portion (steak-bones, peach-pits) quantification
3. Code foods into numbers	Small code grouping of items with loss of information Reading errors Writing errors
4. Portion size conversion to gram amounts	Plate waste Non-standard portions Refuse deduction Portion calculations from standard recipe books
5. Entry of subject identification, date, meal, foods and amounts	Transcription errors
6. Checking entered data for correctness	Oversights Difficult forms No printout of food names
7. Correction of errors	Renewed typing errors
8. Recheck entered corrections	Oversight or elimination altogether
9. Calculation	Non-standard, non-debugged programs

That is an overview of the work involved. Now let's concentrate on the code itself.

APPROACHES TO CODING

The code is some specific and unique identification for each substance in the data base. The code may or may not contain information intrinsically and this choice of an informational or non-informational code helps determine the desired approach to coding.

The code itself may be numerical, alphabetical or mixed. One may have direct access without deliberate coding, that is, no need to enter a code and one may have no code at all. These options are illustrated in Table 2, along with some examples.

TABLE 2

APPROACHES TO CODING

1. Numerical code	123 18607258345
2. Alphabetical code	hamburger zygrhrjztsjr
3. Mixed code	b72
4. Direct access	touch it point to it
5. No code	lists pre-programmed questionning

Numeric codes can be as elegant and simple as 1-2-3, as in the first example.

They may, on the other hand, while trying to pack too much information into the code number (which is seldom necessary) result in a grossly awkward code along with proportionally more errors in transcription.

Alphabetical codes may apply easy-to-remember codes, such as simple food names or acronymes, or more cumbersome ones which for some reason can come out as with the second example; incorporating varied information such as that this is a mixed food, eaten in a restaurant, grilled, from frozen meat, using ground round, top choice pure beef, and certain spices and so on. This example is extreme, but I have seen some difficult to recognize and remember computer generated alphabetic codes. And I can imagine that mistakes are more frequent with such alphabetic codes than number codes because letters are not as "neutral"; one wants to associate foods with particular letters or sequences. Alphabetical codes of equal character length with numeric codes do require less storage space since 26 possibilities, rather than 10, are packed into each 2 bytes. An important drawback of alphabetical codes is that the coder must be able to type--that is to find the letters on the keyboard, rapidly and accurately; which is no minor consideration.

Mixed codes, a combination of letter and numbers, may improve coding--if, for example, the first letter is the food group (as in the example in Table 2, b for beverage, and then the number 72 for its specific identification. This can keep the numbers low, if designed sensibly.

Direct access incorporates newer technology such as touching the desired code on a screen, rather than searching for it in a book and transcribing it. Another approach to direct access, is, as we code in Heidelberg pointing to the desired food on the screen with an arrow--which will be described. The final option in Table 2 is no code--which sound very attractive; a system in which the coding is done automatically. This requires rather advanced technology such a pre-programmed 24 hour recall directly on the computer or list of foods in which the subject himself makes the match in searching for his appropriate choice. We have done the former with the 24 hour recall (2) and are quite pleased with the results.

NUMERIC CODING SYSTEM

Most people use numeric codes, and the numbering system can take various forms, as alluded to earlier. The code number may be random; absolutely random and totally non-informative. Or it can be an alphabetical code so that apple is the number one and zucchini is last (which is difficult to update), or sequentially numbered on the basis of entry into the data base, so that each new food in the data base--and they do keep expanding--gets the next number. With this one can easily tell when what came into the system, but one cannot easily find the codes.

The numbering system may be partially hierarchical; the first few digits for a food group, and then a sequential numbering within a food group for example, perhaps ordered after frequency of use. Another alternative, more difficult to develop, is a strictly hierarchical code in which every digit has a meaning, and the number of digits is determined by the required exactness of description; up to species, brand type of code with which we are operating. It is called the Bundeslebens-mittelschlussel (BLS). This is a "Federal Food Code", now contains approximately 5000 foods and 3000 mixed dishes, is 3 to 12 characters long, and strictly hierarchical in form, as illustrated in Table 3, along with an example (3).

TABLE 3

STRUCTURE OF THE BLS

Level		Example	
1-3	Basic food	2	Vegetable
4	Industrial processing	26	Root vegetable
5	Preparation method	263	Carrot
6,7	Modifications of basic foods	2636	Cooked carrot
8	Source, where purchased		
9	Packaging		
10-12	Country of origin		and so on.

The code is expandable and collapsable to the desired number of digits so that if one is not interested in positions 6-12, coding with 5 digits, as shown here, can be carried out. One may also drop earlier digits and code for example levels 1-3 and 12 as a 4 digit code (Rottka et al.).

A final approach to numeric codes incorporates "safety digits". The last digit of the code is related by complex mathematical equation to the first digits so that a one digit transcription mistake can be readily recognized, and an invalid code number results--the beginning and end do not match up properly. It is a method of error recognition but requires an increase in the number of digits in the code. In my opinion, simpler is better; as most people want to pack a maximal amount of information into a microcomputer; and other methods of entry without the risk of transcription error are available.

ALPHABETICAL CODING SYSTEMS

Another option is to use alphabetical codes. This presents problems, even if food names are used, as multiple names are possible for a single substance; such a frankfurter, hot dog, sausage, and bratwurst; and therefore duplicate listings are necessary, at the expense of storage space.

It is name descriptions in general which make the greatest demands on storage space. We allow 24 characters per name, which is like having a 24 digit code, or equivalent to sacrificing the storage of 12 nutrients with that food.

Also to be considered when using alphabetical name codes is the problem of syntax. Exact syntax is required by all small programs, and the acceptable syntax must be thoroughly learned by coders or multiple listings again incorporated, which can really tax the storage limits of the microsystems when you have a basic data base with a few thousand foods. Table 4 shows a few examples with milk.

TABLE 4

SYNTAX COMPLICATIONS

Milk, full fat
Full milk
Cow's milk, full fat
Full fat milk
Dairy products, milk...

Potential syntax problems can be tackled with a mixed system. The coder uses the basic, common food name or food group name and then chooses the specific or qualified foodstuff from a list of options.

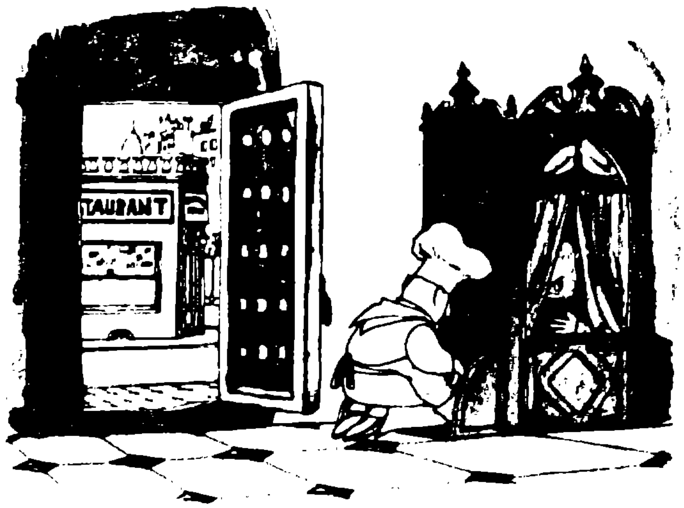
INVISIBLE CODING SYSTEMS

If not calculated by hand, all foods must be stored and identifiable through some system by the machine. There are methods in which the coder is not troubled with this process. For example, presentation by lists which can be scrolled up or down on a screen can be made to the coder. Another possibility is a logical questionning based on a decision tree with many possible options which automatically presents the foods. Or, a hierarchical code with level by level access to the desired food and preparation over food groups can be developed on the computers so that the coder need only point to the food group or specification of interest and thereby come closer and closer to the desired object. This is only programmable in a logically developed code and is the way in which coding is done at our institute (4). In all three methods the code number is invisible to the coders.

PROBLEMS IN CODING

The problems in coding mentioned include syntax variability, transcription errors, and multiple names for the same item. A greater problem is the use of the same name for two completely different items. An example of this, shared with me by Dr. Ken Samonds, is the recalling of Boston school children of eating multiple milk shakes each day. Closer questionning revealed that a candy bar named Milkshake was the real food consumed. One must be careful in the data collection stage to minimize such phenomena--one must be familiar with the population's behavior and its food choices. Fantasy names are also difficult to code--such as the name of one of my favorite Chinese dishes "ants climbing a tree"; or things like "sunshine salad".

Then there is the problem of mixed dishes and prepared dishes. Coding a simple food such as avocado is basically straightforward although the fat composition will differ between those from California and those from Florida--but it's the least of many problems. How does one code something like a salad? Numerous items and their proportions need to be identified and coded (unless your code has extensive mixed dish recipes), each item separately. Mixed dishes present other problems, especially if eaten away from home; one being that the consumer does not know exactly what he ate--only the chef does, as illustrated in Figure 2 (5).



Another problem is that of yield and retention factors. A cooked mixed dish, such as a casserole, cannot accurately be coded by its raw ingredients, that is not what is consumed. Yield and retention factors after cooking are relevant. An ideal code does this for you--has many standard recipes, brand names and options for coding various eaten forms of the basic food: peeled/unpeeled, cooked/raw, canned, dried, etc., so that a meal as in Table 5 can be coded with 6 codes rather than 26.

TABLE 5

SAMPLE MEAL WITH MIXED DISHES AS CODED WITH BLS

Amount	Food	Code
200 g	Greek salad	8301
1 slice	Mushroom pizza	8713
250 ml	Beer, Pils	0126
1 med.	Apple without peel	1182
2 scoops	Ice cream, vanilla	0345
3 Tbs.	Hot raspberries	1323

AMOUNT ESTIMATION

A huge problem which will not be elaborated on is that of amounts eaten. This is probably the greatest source of error in dietary assessment. Unfortunately amounts must also be entered, but the accuracy of estimated amounts is notoriously bad. Using ranges would probably be much more sensible. Measuring the recipe content of a meal is relatively simple compared to assess what people really consume. Depending upon the design and purpose of the exercise, different approaches can be taken. One can provide scales and ask people to weigh everything, but this is expensive, difficult for the subject, complicated to do properly in real life and probably has a great effect on what they eat. One can ask subjects to estimate amounts, which we have found to be unilaterally inaccurate. One can use portion size models--either realistic or abstract. Finally, many researchers depend on some normal serving size standards, which is presented as an option in our direct coding-on-computer system. When buying a nutrient analysis system it would be worth asking if household units and standard portion sizes are built in. What most people will want to enter into their data bases or calculations is person identification, meal food codes and amounts.

ENTERING PROCEDURES

Entry options for food codes and amounts include the old standby, rapidly fading away now, and for good reason, punch cards; optical readers using specially developed forms that need to be filled out specially in specific areas; and direct entry via computer terminal. In Heidelberg the latter is applied to these tasks.

As mentioned under the topic invisible codes, a computerized code book, based on a hierarchical code (the BLS) which requires no typing skills and no transcription of numbers in coding is used (4). Code and entry occur jointly with immediate checking and documentation. The program begins with asking the subject, date and meal--and if you don't know the code, one can opt to find the food by food group. Because the code is hierarchical one can find food within the 4-5 levels. 1) food group; 2) subgroup; 3) exact item; 4) how eaten; 5) how stored. The person at the computer need only use 4 keys on the keyboard to find a food. Those keys place an arrow up or down and to a higher or lower level of code. Then amounts are entered--next to the food name, and pressing return allows entry of the next food. A correction mode is possible at the end of any entry prior to storage or calculations. A final printout of foods and amounts in the same format as the original dietary records allows documentation of input and simplifies cross-checking.

Another approach which we apply is direct entry of foods, bypassing the coding stage, with the previously mentioned 24 hour recall on computer. In this system, the computer asks what was eaten for breakfast for example, and offers a choice of several food groups, selections here lead to a more detailed probing--and the desired description is stored by the computer without any steps of "pen to paper". This quantified information is sent to floppy disks and stored there until calculation or transfer to the data base. This questionning, coding, entry and storage occur simultaneously--eliminating many of the potential errors listed in Table 1.

CONCLUSION

Finally, to summarize, coding and entry are important issues in nutrient analyses, and there are numerous possible ways to do this. The best choice depends on the number of foods in the data base, the nutrients of interest, food groups of interest, ease of location of a food, ease and accuracy of entry, storage space constraints and access time.

Unfortunately, the best code for ease of use by coders is often not the most desirable to the programmer and may or may not meet the needs of research orientations. Coders want direct access, preferably no writing and little searching. If a code must be read and written, then it should be short and simple--acronym letters and lists for those who can type; numbers for those who can't. Programming is easiest and access fastest with hierarchical codes on a base 10 system. Researchers often want to be able to group foods, and therefore have some hierarchy in their code. They also want to know total amounts of basic foods consumed, and would therefore prefer not storing information as mixed dishes.* Priorities need to be set, and we set ours with a coder, who in the authorities opinion have the most tedious job, who are most responsible for data quality, and who spend the greatest amount of time with the task. An alternative option is to have two cross-referenced codes and a conversion program, or to use invisible codes with the type of direct entry systems we are applying.

My parting remark is that a standard food code available for use by all groups would improve comparability of results and facilitate recalculation of nutrient content as new data and new tapes arise. I believe it would prove to be well worth the effort.

*It is very difficult to add up the number of eggs used for example, when hidden in breads, noodles, casseroles, etc.

REFERENCES

1. Arab, L. Schellenberg, B., Schlierf, G. (1982): Nutrition and Health--a survey of young men and women in Heidelberg. In: Ann. Nutri. Metabl 25/S1/82. Karger, Basel

2. Arab, L., Bellin, O., Schlierf, G. (1983): Ein standardisiertes System für das 24-Stunden-Ernährungs und Activityätsprotokoll, das die Befrager-Variabilität ausschließt. XX. Wiss. DGE-Kongreß Gießen. Ernährungsumschau 30/7, S. 249.

3. Rotte, H., Arab, L., Polensky, W. (1983): Bundeslebensmittelschlüssel (BLS) für Verzehrserhebungen. XX. Wiss. DGE-Kongreß, Gießen. Ernährungsumschau 30/7, S. 250.

4. Arab, L., Pfannendörfer, H., Schlierf, G. (1983): Lebensmittel codieren ohne Nummern--ein Computer programm zur direkten Dateneingabe. XX. Wiss. DGE-Kongreß, Gießen. Ernährungsumschau 30/7, S. 249.

5. Quino: Cartoons. Deutscher Taschenbuch Verlag, München, 2. Aufl. 1980.

THE PROBLEMS OF USING SMALL NUTRIENT DATA BASES

Loretta W. Hoover
Professor
Department of Human Nutrition, Foods,
and Food Systems Management
College of Home Economics
University of Missouri-Columbia

Determining what size of nutrient data base to use is a decision which many face when planning to use computers. Reduced nutrient data bases are of two basic types: (1) grouped data consisting of weighted averages for nutrients used to represent all foods in the group and (2) direct substitution where the nutrients for one food are substituted for the nutrients in a somewhat similar food. Pennington (1) has described a procedure for establishing a "mini-list" by using correlation to determine the co-occurrence of seven index nutrients in foods. Investigators have demonstrated that both of these types of small nutrient data bases were functional for their intended purposes.

Major studies such as the USDA Nationwide Food Consumption Survey (NFCS) (2) and the HHS National Health and Nutrition Examination Survey (NHANES) (3) have identified foods commonly eaten by the U.S. population. These foods have served as the basis for smaller data bases.

What distortions in nutrient values can one expect when fewer foods are included in a nutrient data base? In response to that question, the findings from a study involving the use of two reduced nutrient data bases are summarized in the first segment of this paper. The final portion of the paper is focused on my personal opinions and concerns relative to some potential problems associated with the use of small nutrient data bases.

COMPARISON OF LARGE AND SMALL NUTRIENT DATA BASES

With funding from USDA*, a project was conducted at the University of Missouri-Columbia to examine the effect of food item specificity on mean nutrient intake values reported for different sex/age groups using data from the first quarter of the 1977-78 Nationwide Food Consumption Survey of Individuals. The project consisted of four phases: (1) loading the USDA for use on the UMC computer, (2) analysis of food consumption frequencies, (3) developing a cross-reference file for substitutions, and (4) reducing the nutrient data base and testing the results for two substitution levels. The nutrient data base contained values for energy and fourteen nutrients for 100 gram portions of 4,404 food items. The individual intake data processed in the study were three-day dietary records for 7,914 individuals grouped into 22 sex/age groups.

* Funding for this research was provided by the U.S. Department of Agriculture under Research Agreement Number 53-32U4-0-201.

Analysis of Frequency of Consumption

During the second phase of the project, the frequencies of consumption were determined for each food item. Several types of output were generated to list the frequencies in several formats: (1) descending order of frequency, (2) in order as maintained in the nutrient data base and (3) in descending order in each major subgroup. Items eaten several times a day had higher frequencies. The consumption frequencies were displayed according to sex/age group in the third format. The total frequencies for the 4,404 foods were tabulated into three categories: 1371 foods had zero frequency, 2574 foods were consumed 1 to 99 times, and 459 foods had frequencies greater than 100. Thus, over 25% of the food items were not consumed during the first quarter of the survey. The relationship between rank order and frequency was established with a rank of 1 assigned to whole fluid cow's milk with a consumption frequency of 14,142. A reciprocal relationship existed since the consumption was ranked in descending order; a food with a consumption frequency of 88 was assigned a rank of 500.

Nutrient Data Base Reduction

Prior to reduction of the nutrient data base, the food items were analyzed for similarity in nutrient profile. A total of 1044 food items had nutrient values identical to another item in the nutrient data base and differed only in food item description. To reduce the nutrient data base, frequently consumed foods which were unique in nutrient composition and could be representative of a group of similar foods were designated as Retained Items. Food items were designated as Inactive by grouping with a Retained Item assigned as a substitute in a cross-reference file which was created to maintain the substitution assignments. The nutrient data base was not altered. The clusters of Retained and Inactive items were analyzed for similarity in nutrient profile. In Substitution Level 2, the goodness-of-fit between nutrient profiles for the Retained and Inactive items were analyzed using regression analysis. Greatest emphasis was placed on achieving a better fit, for the most frequently consumed foods with a rank greater than 500. R^2 values for 80% of those 499 foods ranged between .9 and 1.0 and the Beta values ranged between .7 and 1.3 for 83% of those foods indicating satisfactory substitution assignments for most of the frequently consumed foods. Although a similar analysis was not made for Substitution Level 1, the fit was assumed to be better since more foods were retained.

Substantial reductions were made in the size of the nutrient data base. In Substitution Level 1, 396 food were retained resulting in a 91% reduction. In Substitution Level 2, a 95.4% reduction was achieved by reducing the nutrient data base to 200 foods. The effective reduction was really somewhat less since some foods were not consumed and some foods had nutrient values identical to another food maintained as a Retained Item. The number of foods retained in the two substitutions levels for each major subgroup are shown in Table 1.

Consequences of Nutrient Data Base Reduction

The nutrient values per sex/age group were computed for the total sample using the original nutrient data base to serve as the baseline for comparison with the values generated using the reduced nutrient data bases. This analysis was also performed to assure that the computer program was accurate. The nutrient values were computed for both substitution levels and were compared with the baseline nutrient values.

Percent differences from the baseline values were computed for each day and the three-day average for each nutrient for each of the 22 sex/age groups. The minimum and maximum absolute percent differences were determined for each nutrient for both substitution levels. The maximum values were larger in Substitution Level 2 and many of the larger values were attributed to the Under 1 age children since very few baby foods were retained. For Substitution Level 1, the maximum values were under 10% for all nutrients except magnesium and Vitamin A. In Substitution Level 2, Vitamin A and Vitamin B12 were associated with percent differences greater than 10% after eliminating the differences attributed to the Under 1 age group.

A two factor analysis of variance model was processed for each nutrient for both substitution levels to statistically analyze the effect of nutrient data base reduction. In Substitution Level 1, F-ratios were significant at the .05 level for fat, carbohydrate, calcium, iron, magnesium, Vitamin A, thiamin, and Vitamin B6. In Substitution Level 2, only Vitamin B6 was not associated with a significant F-ratio for either data base type or interaction between data base type and sex/age group. Thus, the results for Substitution Level 2 were more complex and difficult to interpret. A statistical analysis of the power of the F-test was performed to determine minimum detectable differences in mean nutrient values.

The consequences of the reductions were also analyzed by performing t-tests for each nutrient for each of the 22 sex/age groups. A total of 330 t-tests were performed for each substitution level; 199 were significant at the .05 level in Substitution Level 1 and 207 were significant in Substitution Level 2.

Since nutritional adequacy is often expressed in terms of a standard such as the Recommended Dietary Allowances (RDA), that standard as adapted by USDA staff (2) was used for comparison with the baseline nutrient values and those from Substitution Levels 1 and 2. The comparisons were made by computing percent differences using the three-day average values for 13 nutrients. Except for six instances, the differences all carried the same sign. Using the reduced nutrient data base, the conclusions about nutritional adequacy were essentially the same as those made using the original nutrient data base.

Conclusions and Recommended Research

After considering the details of these analyses, several conclusions seemed appropriate. If a nutrient data base is tailored to consumption practices, the nutrient intake of a large group can be approximated with a smaller data base. However, translating those results into nutritional guidance may be difficult since the specific food selections are not known. Small nutrient data bases are probably inadequate for analyzing nutrient intake for specific subsets of the population and may not be suitable for intake data from other quarters of the NFCS. Small nutrient data bases are not suitable for individual dietary records or data from small groups if numerous substitutions must be made. One of the chief reasons for questioning the use of a small nutrient data base is that specific food consumption practices of a diverse population cannot be monitored effectively. Thus, a sound basis for nutritional guidance may not be feasible using small nutrient data bases which require many substitutions.

More research is needed to determine the effects of reduced nutrient data bases on other quarters of data from the NFCS and to determine the effect on nutrient values for subsets of the population. Also, the suitability of a reduced nutrient data base larger than those tested in this project should be evaluated to determine the value of a mid-size nutrient data base.

LIMITATIONS OF SMALL NUTRIENT DATA BASES

Some individuals use small nutrient data bases for a number of reasons. Small nutrient data bases have fewer options to consider when coding, fewer items to keep up-to-date, require less computer data storage and memory, are not cluttered by many rarely eaten foods, and may be tailored to meet the needs of a specific project. Some regard the types of variations identified in the research project described above as within the range of normal variation for nutrients in foods and are not concerned about the distortion introduced by data analysis. Other developers prefer smaller nutrient data bases since the probability of missing values may be reduced. However, the potential problems with small nutrient data bases should be considered.

Potential Problems of Small Nutrient Data Bases

Potential problems may be associated with using small nutrient data bases. Whether or not these problems exist may depend on the size, contents, and intended use of a small nutrient data base. Some of the problems are itemized below:

1. May lack food item specificity.
2. May not represent foods eaten in various seasons.
3. May not contain foods eaten by sex/age groups, geographic areas, or ethnic populations.
4. May not contain a variety of mixed dish items.
5. May not reflect various food preparation methods.
6. May not reflect intake of some nutrients due to lack of specificity resulting in incorrect estimation.
7. May discourage specificity in data collection if fewer food items are available for coding.
8. May require substitutions for numerous foods.
9. May result in more inconsistency in coding food items and poorer intercoder reliability.
10. May require more judgment in coding food item substitutions and result in frustration for coders.
11. May provide inaccurate diagnostic information if numerous substitutions are required resulting in an imprecise screening tool.

12. May bias results through decisions made while coding.

13. May not have a sound basis for dietary guidance.

14. May not satisfactorily reflect nutrients of concern in future studies.

15. May be limited to use in certain studies since not suitable as a general purpose tool.

16. May not be appropriate for analysis of individual dietary records.

17. May not include forms of food needed for calculation of nutrients for recipes.

18. May need custom software to select and update records when new nutrient data are released.

19. May multiply coding effort by coding several items to represent a single mixed dish item.

Summary

Each professional has the responsibility to select a nutrient data base which is appropriate for intended uses. With the advances being made in computer storage technology, size limitations on nutrient data bases used on microcomputers are being eliminated and larger nutrient data bases can be accommodated.

References

1. Pennington, J. A.: Dietary Nutrient Guide. Westport, Conn: AVI Publishing Company, Inc., 1976.

2. Food and Nutrient Intakes of Individuals in 1 Day of the United States, Spring, 1977. USDA Nationwide Food Consumption Survey, 1977-78. Preliminary Report No. 2. Science and Education Administration, Washington, D.C., 1980.

3. Dietary Intake Findings, United States, 1971-1974. Data from the National Health and Nutrition Examination Survey. DHEW Pub. No. (HRA)77-1647, 1977.

TABLE I

NUMBER OF RETAINED FOOD ITEMS IN NUTRIENT DATA BASES

FOOD GROUP	*NUMBER OF FOODS IN ORIGINAL NDB*	*SUBSTITUTION LEVEL 1*	*SUBSTITUTION LEVEL 2*
Milk and Milk Products	321	36	24
Eggs, Mixtures and Substitutes	51	8	3
Fats, Oils, and Salad Dressings	70	11	5
Sugars, Sweets, and Beverages	392	48	14
Grain Products	956	68	47
Dry Legumes, Nuts, and Seeds	157	13	10
Meat, Poultry, Fish, and Mixtures	1307	107	60
Vegetables	677	67	23
Fruits	473	38	14
Total	4404	396	200
Average Percent Reduction	—	91%	95%

PROBLEMS ASSOCIATED WITH RECIPE ANALYSIS

Anne Marsh
Nutrition Analyst
Human Nutrition Information Service
U.S. Department of Agriculture

The Nutrient Data Research Branch at USDA's Human Nutrition Information Service is responsible for the publication of Agriculture Handbook No. 8, "Composition of Foods...Raw, Processed, Prepared." We are in the process of revising the handbook in sections by food group. The section to which I am currently assigned is the one covering mixed dishes. For this handbook section, we need values for the cooked forms of mixed dishes since this is the form in which they are consumed. In trying to derive nutrient values for these dishes, we have been looking at various ways to calculate the values from ingredient information, based on formulas or recipes. Of course, we want to do this work by computer. Dr. Jane Wyatt, at Oregon State University, has undertaken a research project to analyze 25 mixed dishes, raw and cooked. The dishes were prepared from recipes, were weighed and analyzed raw and cooked. We have been using results of this research to test the appropriateness of the various methods for calculating recipes from ingredient information.

There are several methods for calculating the nutrient content of a cooked dish, given the raw recipe. I will discuss three methods we have used, and will compare our calculated values with Dr. Wyatt's analyses. In the first method, we apply a retention factor to each raw ingredient for each nutrient to be determined, and we apply a yield factor to the overall weight of the recipe to account for changes in moisture or fat content during cooking. For this method a computer program has been developed. The program uses the Standard Reference Tape which has been discussed earlier.

The second method uses nutrient data for the cooked form of each ingredient. Retention factors for the nutrient are, therefore, not needed but it requires information about yield of each cooked ingredient from raw material. The third method uses nutrient data for each raw ingredient, as in the first method, but a retention factor for the whole dish (not each ingredient) is used for each nutrient. The retention factors and yield information needed for the calculations of each method are found in--

Ag Handbook No. 102--Food Yields

ARS 62-13--Procedures for Calculating Nutritive Values of Home Prepared Foods--Tables 2 and 22

Provisional Table on Percent Retention of Nutrients in Food Preparation

Because the dishes I have worked with are the casserole type--that is, all the ingredients are in the dish, nothing is drained or discarded--I will limit my discussion to vitamins. Any minerals that might leach out of the ingredients would be retained in the broth and there would not be any destruction of minerals. I will discuss three B-vitamins and vitamins A and C.

The first dish is a beef stew without tomatoes (recipe, Table 1). The NDB number represents the identification number for the Standard Reference Tape. We have already converted the volume measure to weight in grams. The ingredients in order of weight are water, beef, potatoes, onions, carrots, frozen peas, flour, salad oil, salt and pepper. The beef is coated with flour and browned in the oil. Water is added and the meat simmered. The remaining ingredients are added and cooked until tender. In this dish, all raw ingredients are added to the "pot" and the only change after cooking is a loss of moisture. Table 2 shows the vitamin content of beef stew before cooking, as calculated from the nutrient data base.

In the first method of calculation, we use the nutrient content of the raw ingredients and apply a retention factor for each nutrient. These retention factors are found in either ARS 62-13 or the Provisional Table. For the first ingredient, flour, multiplying by the retention values for flours and grains from the Provisional Table (thiamin 80 percent, riboflavin 90 percent, and niacin 90 percent) we get .38, .26, and 3.5 milligrams/100 grams for the content of "cooked" flour (Table 3). As mentioned before, we already have a computer program to calculate a recipe for this method. The retention factors are stored in the program. When we enter the identification code for flour, we enter the appropriate code for this set of retention factors. Since wheat flour contains no vitamin A or C, we can disregard them for this ingredient. The factors we will use for both onions and carrots are those for root vegetables prepared from raw and drained. In our beef stew, the vegetables actually are not drained but this is the best retention figure available to us. If our calculation gives a lower vitamin content than was found on actual analysis, this retention factor might be the cause. For potatoes, we will use factors for potatoes boiled without the skin and drained, although like the root vegetables, they are not drained. The factors for peas will be those for starchy vegetables prepared from frozen and drained. Braised beef seems to be the best match for the beef in our stew, so those factors will be used for the B-vitamins in beef. Since no factor is given for vitamin A retention in beef or pork, the 75 percent retention for chicken meat will be applied.

Now we can add the values to determine the content for the total "cooked" dish. If this is divided by the weight of the cooked dish, we can calculate the amount of each nutrient per 100 grams. Determining the weight of the cooked dish presents another problem. The total weight of the raw dish is 6,649 grams. For beef stew, the only change is a moisture loss during cooking. Using a value of 13 percent for evaporation loss in stew, as suggested in ARS 62-13, will give a weight for the cooked dish of 5,784 grams. The values we have derived for beef stew using this first method of calculation are .069, .071, 1.28 mg for thiamin, riboflavin and niacin, respectively; 1,742 IU for vitamin A, and 6.3 mg for vitamin C per 100 grams (Table 4).

The second method for calculating cooked dishes is to convert the weight of the raw product to the weight we would expect after cooking and apply the nutrient composition of the cooked product. The overall weight change will again be a moisture loss of 13 percent--or 864 grams--with a total weight of 5,784 grams. Since the vegetables remain in the cooking container, we will not calculate a moisture change for them. The stewing beef is browned in oil and the remaining ingredients are added to the cooked beef so there is no dripping loss or fat uptake. There will be a loss of volatiles in the beef and the best estimate for this loss is 17 percent from Food Yields--Item #304, pan fried steaks, boneless. The weight of cooked beef is 1,130 grams, a loss of 232 grams. The remaining moisture loss, 632 grams, we will deduct from water to give 760 grams of water in the "cooked" stew.

If we change these weights and use codes for cooked items, with no retention factors, the computer will generate the nutrient composition per 100 grams for cooked stew. The values derived by our second method are .08, .08, 1.7 mg for thiamin, riboflavin, and niacin, respectively; 1,845 IU for vitamin A, and 7 mg for vitamin C per 100 grams (Table 4). These values are slightly higher than those derived by method one--perhaps because we used retentions for drained vegetables.

The third and last method for calculating "cooked" dishes is to apply the dish retentions to the raw recipe. For example, one retention for each nutrient per dish, instead of a retention for each nutrient in each ingredient. In method one, when we applied retention factors to raw data, we did derive a total for each nutrient in the cooked dish. We have the total for the raw dish and by dividing the cooked total by the raw total we get our "dish retention" (Table 5). A dish retention could have derived by method two also. This third method involves only one calculation for each nutrient in the whole dish, not separate calculations for each ingredient, and if the resulting values seem correct, it would be simpler and could be easily handled by computer. The results for the third method are .069, .072, and 1.28 mg for thiamin, riboflavin, and niacin, respectively; 1,743 IU for vitamin A, and 6.5 mg for vitamin C per 100 grams (Table 4).

These results are very close to Method 1 because the dish retention factors for this paper were directly derived from the results of Method 1. Studies are being conducted to provide a basis for dish retentions. If any uniformity could be found for dishes cooked in a certain way--that is, oven baked or cooked on top of the stove--perhaps a set of factors could be derived for the entire group. Whether length of time in cooking or temperature to which mixtures are cooked is the determining condition has not been resolved. The source of carbohydrate--that is, pasta, legume, or rice--and the pH of the dish may also influence the retention of vitamins on cooking.

To see how well our methods of calculation are predicting the actual content of the cooked dishes, let us look at Dr. Wyatt's laboratory results. For beef stew, the analytical values are thiamin .07, riboflavin .09, niacin 1.3 mg, vitamin A 4,007 IU, and vitamin C 3 mg per 100 grams (Table 4). One discrepancy seems very obvious. The analyzed vitamin A value of over 4,000 IU is very high--the raw stew had a content of 1,605 IU. The laboratory repeated the analyses on both dishes and confirmed the first results. The research is not completed at this point, but the explanation is that the extraction is more efficient when the dish is cooked.

The dish we have been considering so far, beef stew, has only a moisture loss during cooking. An additional problem exists when we "cook" a dish like spaghetti, macaroni and cheese, or beans and franks in which water is absorbed by a raw ingredient. Let us consider spaghetti with tomato sauce (recipe, Table 6). Using the first method--raw ingredients times retention factors--the nutrient content of the dish would not be affected by this water absorption (Table 7). The problem would be in arriving at a weight of cooked dish. We are given a weight of raw spaghetti of 680 grams and using a yield factor of 282 percent, from Food Yields, we get 1,918 grams as the weight of cooked spaghetti.

ARS 62-13 gives an evaporation loss of 10 percent for spaghetti in tomato sauce. Our dish weight with 1,918 grams of cooked spaghetti is 3,344 grams, and a 10 percent loss gives 3,010 grams as the weight of cooked spaghetti in tomato sauce. The nutrient content per 100 grams is .18, .10, and 1.6 mg for thiamin, riboflavin, and niacin, respectively; 809 IU of vitamin A, and 14.4 mg of vitamin C per 100 grams (Table 8). We can derive "dish retentions" as we did for beef stew.

In method 2, we would use the same weight for cooked spaghetti and use the nutrient composition for cooked spaghetti, cooked onion, and cooked peas. Our weight of cooked dish would be the same as in method 1 and the results of this calculation per 100 grams are .14, .09, and 1.5 mg for thiamin, riboflavin, and niacin, respectively; 850 IU of vitamin A and 15.0 mg of vitamin C per 100 grams (Table 8).

Using the "dish retention" the same way we did in the beef stew we get .18, .10, and 1.61 mg for thiamin, riboflavin, and niacin, respectively; 808 IU for vitamin A, and 14.4 mg of vitamin C (Table 8).

The laboratory-analyzed results for spaghetti with tomato sauce are thiamin .13 mg, riboflavin .07 mg, niacin 1.8 mg, vitamin A 704 IU, and vitamin C 4.6 mg (Table 8). The calculated values for vitamin C are higher than the analyzed value. There may actually be destruction in freeze-drying the sample.

A dish with fat uptake on cooking, zucchini fritters (recipe, Table 9), presents additional problems to be considered in calculating the cooked dish. ARS 62-13 gives 35 percent evaporation loss and a 14 percent fat uptake on cooking for corn fritters. We will use these figures for zucchini fritters. The first method with raw ingredients and retention factors is straightforward, but there is a net loss of 21 percent due to evaporation and fat uptake. The weight of the raw dish is 1,754 grams and deducting 21 percent or 368 grams gives a weight of cooked dish of 1,386 grams. The values obtained using this method are .22, .26, and 2.0 mg for thiamin, riboflavin, and niacin, respectively; 268 IU for vitamin A and 5 mg for vitamin C (Table 10).

In using the calculation method 2, we need a yield figure for zucchini. The value given in Food Yields #102 is 85 percent so we have 464 grams of cooked zucchini--a moisture loss of 82 grams. As stated before, we would have a 35 percent total moisture loss but we have already accounted for 82 grams, or 5 percent, of the total, so we can apply an additional 30 percent to the recipe. The 14 percent fat uptake calculation must be done before evaporation and we find 245 grams of fat is absorbed. The values for this calculation are thiamin .26 mg, riboflavin .27 mg, niacin 2.1 mg, vitamin A 252 IU, and vitamin C 3 mg per 100 grams (Table 10). Applying dish retentions gives .22, .26, and 2.0 mg for thiamin, riboflavin, and niacin, respectively; 270 IU for vitamin A, and 5 mg for vitamin C (Table 10).

The analytical results for zucchini fritters are .22, .30, and .86 mg for thiamin, riboflavin, and niacin, respectively; and 349 IU for vitamin A per 100 grams (Table 10). Our calculations for riboflavin are slightly lower than analyzed but niacin appears to be over-estimated. More vitamin A was found on analysis but vitamin C was apparently below the limits of detection.

I have not reached a conclusion as to which method is best. We do have a computer program for method 1 which will apply retention factors. Our program will handle method 2 also, when the identification codes and appropriate yields are inserted for cooked ingredients.

For our purposes, in the mixed dish section, a "dish retention" calculation would be ideal. Much of the data that comes from manufacturers is for uncooked dishes. We have no recipe with which to do calculations, so the total dish retention would be helpful.

Obviously, more research is needed before we can determine "dish retentions" which can be applied to a group of dishes. Our 25 dishes have given us an overview of the problems involved and a project is currently underway to study in more detail the conditions that affect retention.

TABLE 1.--BEEF STEW RECIPE

Standard reference name	NDB no.	Weight/grams
Wheat flour--------	94390	74.0
Salt---------------	89630	19.8
Pepper-------------	2030	1.2
Beef, round, raw---	73520	1,362.0
Oil, soybean-------	4034	51.7
Water--------------	97010	1,392.0
Onion, yellow, raw-	84121	1,048.0
Potato, raw--------	87851	1,268.0
Carrots, raw-------	76190	984.0
Peas, frozen-------	85291	448.0
Total weight		6,649.0

TABLE 2.--VITAMIN CONTENT OF BEEF STEW BEFORE COOKING

Ingredient	Thiamin	Riboflavin	Niacin	Vitamin A	Vitamin C
	Mg	Mg	Mg	IU	Mg
Flour--------	.47	.29	3.9	0	0
Salt---------	.00	.00	.0	0	0
Pepper-------	.00	.00	.0	2	0
Beef---------	1.18	2.45	66.1	341	0
Oil----------	.00	.00	.0	0	0
Onions-------	.31	.42	2.1	419	105
Potatoes-----	1.27	.51	19.0	0	254
Carrots------	.59	.49	5.9	108,240	79
Peas---------	1.43	.45	9.0	3,046	85
Total	5.25	4.61	106.0	112,048	523

TABLE 3.--WHEAT FLOUR (per 74 grams)

Nutrient	Raw value		Retention		Cooked value
Thiamin-----Mg--	.47	X	.80	=	.38
Riboflavin--Mg--	.29	X	.90	=	.26
Niacin------Mg--	3.9	X	.90	=	3.5
Vitamin A---IU--	0		--		0
Vitamin C---Mg--	0		--		0

TABLE 4.--VITAMIN CONTENT OF COOKED BEEF STEW (per 100 grams)

Nutrient	Calculated by--			
	Method 1	Method 2	Method 3	Analyzed
Thiamin-----Mg----	.069	.08	.069	.07
Riboflavin--Mg----	.071	.08	.072	.09
Niacin------Mg----	1.28	1.7	1.28	1.3
Vitamin A---IU----	1,742	1,845	1,743	4,007
Vitamin C---Mg----	6.3	7	6.5	3

TABLE 5.--DISH RETENTIONS

Nutrient	Beef stew	Spaghetti	Zucchini fritters
	%	%	%
Thiamin-----	76	71	81
Riboflavin--	89	82	93
Niacin------	70	76	91
Vitamin A---	90	95	96
Vitamin C---	73	95	70

TABLE 6.--SPAGHETTI WITH TOMATO SAUCE RECIPE

Standard reference name	NDB No.	Weight/grams
Onion, yellow-------------	84121	100.0
Garlic, cloves------------	80290	3.2
Oregano-------------------	2020	.4
Salt----------------------	89630	10.0
Pepper--------------------	2030	1.5
Tomato paste, canned------	92952	340.0
Tomato puree, canned------	92960	850.0
Sugar, granulated---------	92300	20.8
Grated parmesan-----------	1032	100.0
Spaghetti, enriched, dry--	91570	680.0
Total weight		2,105.9

TABLE 7.--B-VITAMIN CONTENT OF SPAGHETTI WITH TOMATO SAUCE

Ingredient	Weight in recipe	Content in recipe		Retention factor		Content retained
	g	Mg				Mg
				Thiamin		
Onion------------	100.0	.03	X	.85	=	.03
Garlic-----------	3.2	.01	X	1.00	=	.01
Tomato paste-----------	340.0	.68	X	.95	=	.65
Tomato puree-----------	850.0	.76	X	.95	=	.72
Parmesan---------	100.0	.04	X	.90	=	.04
Spaghetti--------	680.0	5.98	X	.65	=	3.89
Other ingredients	32.7	--		--		--
	2,106.	7.5				5.34

TABLE 8.--VITAMIN CONTENT OF COOKED SPAGHETTI WITH TOMATO SAUCE (per 100 grams)

Nutrient	Calculated by-- Method 1	Method 2	Method 3	Analyzed
Thiamin-----Mg----	.18	.14	.18	.13
Riboflavin--Mg----	.10	.09	.10	.07
Niacin------Mg----	1.60	1.50	1.61	1.8
Vitamin A---IU----	809	850	808	704
Vitamin C---Mg----	14.4	15.0	14.4	4.6

TABLE 9.--ZUCCHINI FRITTER RECIPE

Standard reference number	NDB No.	Weight/grams
Wheat flour-------	94390	474.0
Baking powder-----	71320	12.0
Salt--------------	89630	19.5
Egg---------------	1123	324.0
Milk, 3.3% fat----	1077	366.9
Vegetable oil-----	84017	12.0
Zucchini, raw-----	91970	546.0
Total weight		1,754.4

TABLE 10.--VITAMIN CONTENT OF COOKED ZUCCHINI FRITTERS (per 100 grams)

Nutrient	Calculated by-- Method 1	Method 2	Method 3	Analyzed
Thiamin-----Mg----	.22	.26	.22	.22
Riboflavin--Mg----	.26	.27	.26	.30
Niacin------Mg----	2.0	2.1	2.0	.86
Vitamin A---IU----	268	252	270	349
Vitamin C---Mg----	5	3	5	*

HOSPITAL APPLICATIONS OF A NUTRIENT DATA BASE

Ann T. Sandvick, R.D.
Manager of Nutrition Services
Department of Food Service & Nutrition
Children's Hospital Medical Center of Akron, Ohio

Children's Hospital Medical Center of Akron is a 253 bed acute care teaching facility providing the full spectrum of services to the pediatric age group, and providing regional services in both neonatal intensive care and burn care.

Approximately 4 years ago, the Department of Food Service and Nutrition of CHMC purchased the HVH-CWRU Nutrient Data Base for use in patient nutritional care. We purchase updates of the data from Case Western Reserve University annually, but have developed our own software due to language incompatibility between the two systems.

DATA BASE SELECTION

The data base was selected primarily because of the flexibility built into the coding system. We utilize the NDB for both inpatients and ambulatory clients. Particularly with diet diary analysis, the availability of flexibility in measuring units allowed for data entry is critical. The HVH-CWRU coding system has arranged foods into 43 groups and all entries are made in the universal 100 gram portion, but with conversion factors that allow the entry to be made in at least 7 and up to 16 household measures or metric equivalents.

INITIAL DIFFICULTIES

The most critical problem was the incompatibility of our computer language, RPG3. CWRU Department of Biometry offered aid in transcription of their programming, but we decided to develop our own.

The other major problem was the need for programming changes from the batch-mode used at CWRU to accommodate online input of data. The first programmer involved had little conception of our needs and failed to make these program changes. Fortunately a new programmer-analyst arrived at this time. Our department also adopted a comprehensive plan for the materials management aspect of The Food Service area incorporating the NDB as an integral part. The new analyst had a background in food service management and time was allotted for him to develop improved programming with an eye to future uses we will make of the NDB.

USES OF THE NDB

Patient Care

There are special needs in our young patient population. Much individualized diet therapy must be provided, since nutritional support of hospitalized children must accommodate their high nutritional needs at a traumatic time of separation from home. We do a great deal of nutrient intake evaluation. The NDB system provides tremendous time savings and leaves the Registered Dietitian more time for medical and patient interaction. We are also able to provide a far greater scope of analysis than hand calculation could ever provide.

Teaching

We also use the NDB system in teaching dietetic and medical students. We are beginning to explore use of the computer in patient education, since the patients who have tested this have been very enthusiastic.

Ambulatory Services

Nutritional assessment includes diet diary analysis at the Physical Fitness Center of a nearby hospital, where the dietetic services are contracted from Children's.

Research

We have submitted 4 grants in 3 years from CHMC of Akron with a nutrition component that incorporated use of the NDB system.

Other Uses

We analyzed all diets in our newly revised 400 page Handbook of Diet Therapies using the NDB. We include a statement of adequacy on each diet. We also analyze all house menus for adequacy whenever our cycle is changed. A popular use has been the posting of kcalorie content of cafeteria food items.

STEPS IN THE DEVELOPMENT OF THE PATIENT CARE USES

The master handbook of all items in the data base was used to code all the items on our 21 day cycle for both patients' and cafeteria items. Thus, we created an efficient and handy "mini code book" personalized for CHMC. We also included all standard write-ins, special dietary products, therapeutic nutritional supplements such as tube feedings and all items available in our nourishment centers. We have then been able to analyze all the items for kcalories, grams of protein, fat and carbohydrate and milligrams of sodium.

Obviously, the subset of an entire code book provides much faster accessibility of the data base. This is essential to make the NDB appealing to the busy professional. This code book is invaluable in the writing of modified diets, and helps allow tremendous variety. Other nutrients are easily available to the dietitian upon request.

One particular problem for a Children's Hospital is the lack of a single standardized portion size. We must offer toddler, child, and adult portion sizes. When we have these worked out we will produce a new mini code book which will reflect these variations in portion size.

Another development needed was efficient coding forms. With the mini code book and these forms, our technician can code and input a patient intake in about 5 minutes. We have coding forms designed for inpatient use and for diet diaries.

Our last critical need was for program outputs which included the summarization of patient intake with a comparison to a standard (the RDA). This form is suitable for inclusion in the Medical Record and is entitled "Nutrient Intake Evaluation."

SPECIAL CONSIDERATIONS FOR THE NUTRIENT INTAKE EVALUATION FORM (attached) In developing the programming for this output we incorporated:

1. Meal by meal analysis. We assume that sometime in the future some nutrients availability (such as Fe) will need to be assessed by meal.

2. Averaging of nutrient intakes. This increases the applicability of the RDA for use in evaluating individual dietary intakes. It is desirable to have 5 days intake.

3. Flagging of incomplete data. This is noted with an asterisk beside the total amount reported. Knowledge of the nutrient content of all foods is far from complete, though growing continually.

Interpretation of these analysis is a critical component for which the dietitian is uniquely qualified. The dietitian must assess foods consumed against nutrient intake, reported particularly when the data is flagged and intake is below two thirds of the Recommended Dietary Allowances (RDA). Other considerations must include the patients health status, medications and other therapies. Without a NDB we could never approach the scope of evaluation we can now provide.

The Materials Management Area

As shown in the schematic, our goal is to have a fully integrated computer system for materials management, with the NDB system at the core. The NDB will be an integrator among food service tasks and between the food service system and the patient care system.

Our Recipe File

The data base has a food group category reserved for recipes. We can use this recipe file to enter "recipes" for vitamin and mineral supplements since individual nutrients are now items in the data base.

We are also working on standardization of recipes as an aspect of materials management.

Future Directions

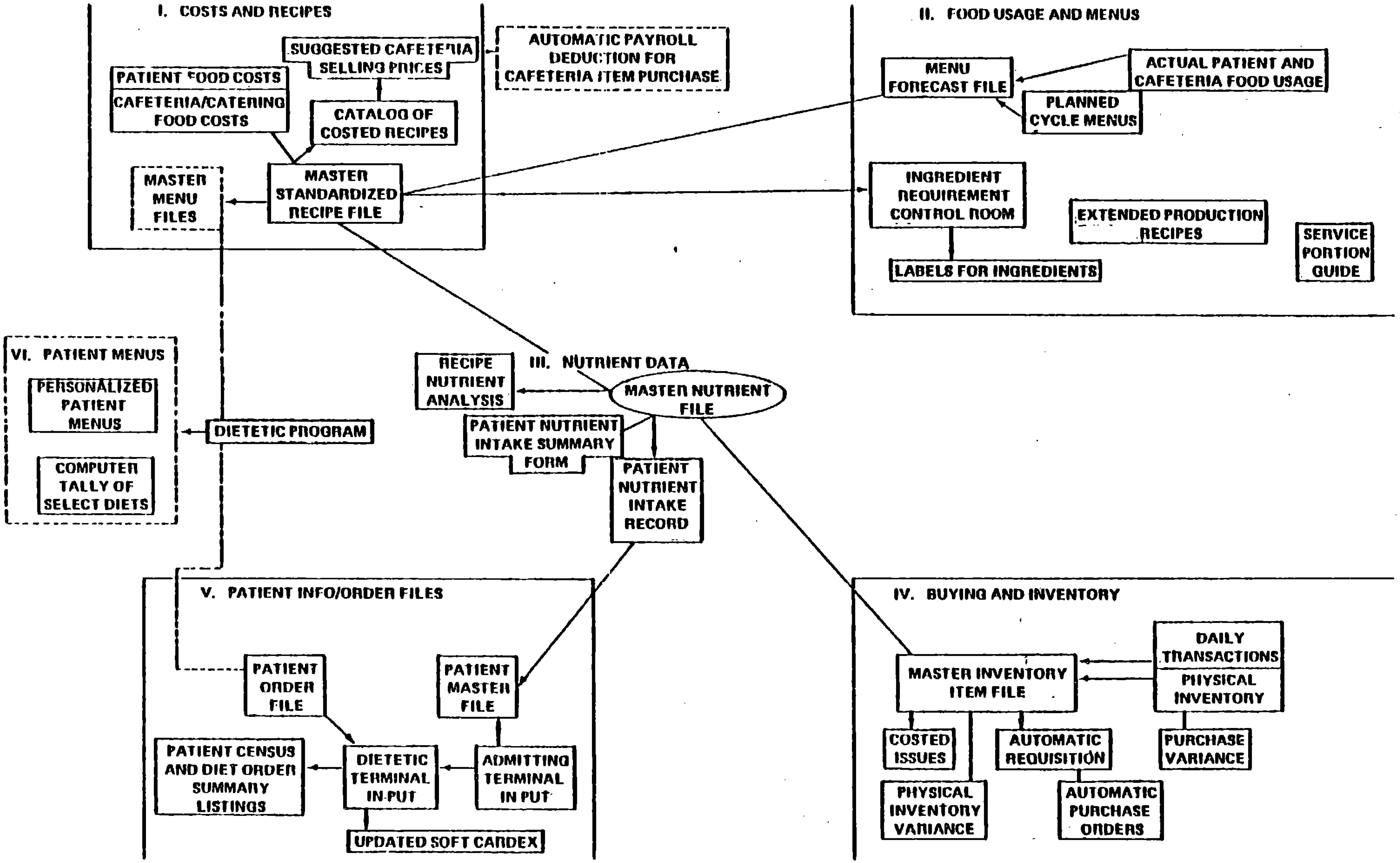
At this time we can summarize all nutrient intake from enteral sources (food, supplements, formulas or tube feedings). We will soon have the capacity to add parenteral fluids, thus we could summarize total nutrient intake from ALL sources!

We are now working on an alternative set of standards to replace the RDA in the evaluating of parenteral nutrient intake.

The Equipment Used at Children's

The hospital system is the IBM 38, which is designed to run efficiently in an online environment. This system has 2 billion bytes of disc storage and an internal memory of 3 million bytes. We use a business communication language RPG3. Our CRT workstation is an IBM Model 5251, and our remote printer is IBM 5256, which has a printing capacity of 300 characters per minute.

MASTER PLAN FOR COMPUTERIZATION OF THE FOOD SERVICE AND NUTRITION DEPARTMENT



1/80
3/82
Children's Hospital Medical Center
Akron, Ohio 44308

CHILDREN'S HOSPITAL MEDICAL CENTER OF AKRON

NUTRITIONAL INTAKE EVALUATION FORM

1 DAY AVERAGE DATE- 07/22/83

PATIENT'S NAME- JOHNNY JONES

			BASED ON RDA FOR AGE (AVERAGE)	FROM NCHS GROWTH CHARTS
NUMBER- 09090909				
HEIGHT- 106.0	CM	41.8 IN	PATIENT IS 88.3 % OF AVERAGE HEIGHT.	PATIENT'S HT/AGE: <50 %
WEIGHT- 15.40	KG	33.9 LBS	PATIENT IS 77.0 % OF AVERAGE WEIGHT.	PATIENT'S WT/AGE: <25 %
AGE—— 4.75	YR		PATIENT'S IDEAL WEIGHT IS 16.5 KG 36.3 LBS	PATIENT'S WT/HT: <25 %
SEX—— M			BASED ON WEIGHT FOR HEIGHT	

NUTRIENT	UNIT	INTAKE AVERAGE	R.D.A. VALUE	PERCENT OF RDA &&	NUTRIENT	UNIT	INTAKE AVERAGE	R.D.A. VALUE	PERCENT OF RDA &&
ENERGY	KC	960.96	1700.000	56.53 %	THIAMIN	MG	1.44	.900	159.57 %
PROTEIN	G	33.60	30.000	111.99 %	VITAMIN B6	UG	1921.92	1300.000	147.84 %
VITAMIN A	IU	2402.40	2500.000	96.10 %	VITAMIN B12	UG	5.76	2.500	230.21 %
VITAMIN D	IU	191.98	400.000	48.00 %	CALCIUM	MG	479.95	800.000	59.99 %
VITAMIN E	MG	28.80	6.044	476.46 %	PHOSPHOROUS	MG	479.95	800.000	59.99 %
ASCORBIC ACID	MG	143.99	45.000	319.97 %	IODINE	MG	.07	.090	80.00 %
FOLACIN	MG	.19	.200	95.00 %	IRON	MG	8.64	10.000	86.38 %
NIACIN	MG	19.20	11.000	174.53 %	MAGNESIUM	MG	191.98	200.000	95.99 %
RIBOFLAVIN	MG	1.63	1.000	163.15 %	ZINC	MG	14.40	10.000	143.99 %

****

NUTRIENT	UNIT	INTAKE AVERAGE	STANDARD
CARBOHYDRATE	G	119.96	50-55 % OF KC
FAT	G	38.40	30 % OF KC
ALCOHOL	G	.00	________
CHOLESTEROL &	MG	.00	< 300 MG PER DAY
POLYUNSAT FA &	MG	.00*	________
SATURATED FA &	MG	.00*	________
SODIUM &&	MG	479.95	450-1350 mg/day
POTASSIUM &	MG	599.81	775-2325 mg/day
ENERGY	KC	960.96	1300 - 2300

CALORIE SUMMARY

CARBOHYDRATE	= 50.06 % OF CALORIES
FAT	= 35.96 % OF CALORIES
ALCOHOL	= .00 % OF CALORIES
PROTEIN	= 13.98 % OF CALORIES

* POSSIBLY AN ADDITIONAL AMOUNT.
& BASED ON 1980 U.S. DIETARY GOALS.
&& BASED ON 1980 RECOMMENDED DIETARY ALLOWANCES.

SIGNATURE Staff R.D.

S: Pt. admitted to ICU for multiple trauma, fractures.

O: Adm. Wt: 15.4 kg (10%/age, 15%/ht) Adm. ht: 106 cm (30%/age) Current Wt: 14.1 kg (1 wk. post adm.) Wt. loss of 1.3 kg since adm. Ideal wt: 17.4 kg (Based on wt for ht @ 50%)

Current nutr. support:

	Protein	Kcals
Osmolyte @ 60 cc q 3 hr	17.6	502
D_5 @ 30 cc/hr		122
TOTAL	17.6	624

Estimated nutr. requirement:

Kcals: BMR x 1.5 = 1087 + 25% for metabolic stress/fractures

725 kcal x 1.5 = 1087 + 25% = 1358 kcal/day

Protein: 30 gm/day (based on RDA for age)

A. Current Nutr. support provides inadequate Kca protein, vitamins, minerals, (except Vit E, V C, Niacin, Riboflavin, Thiamin, B_6, B_{12}, Zinc to meet estimated requirements. Pt. requires approx. 1280 cc Osmolyte/day (160 cc q 3 hr). This amt. of tubefeed will meet or exceed nut needs according to RDA for age & estimated kc needs based on wt. and increased metabolic needs for stress. Potassium intake will meet lower end of recommended range when intake reaches 620 cc or approx. 80 cc q 3 hr.

P: (1) Suggest gradual increase in amt. of tubefeed according to pt. tolerance until desired volume is achieved.
(2) Will post calorie count daily until wt is stable, then 3x weekly. Monitor intake, wt, and tolerance.
(3) Weekly follow up note to update nutrition status, discuss any current problems, etc.

ABSTRACT

"THE NUTRIENT DATABASE: LABELLING AND PRODUCT DEVELOPMENT*

N. Rawson and C. S. Khoo
Campbell Institute for Research and Technology

The use of a nutrient databank in the food industry has gone far beyond the monitoring and labelling functions. Americans now consume more than half their diet as processed foods rather than fresh produce. The fact that the incidence of obesity among Americans is on the increase, contrasted by a decreasing trend in total caloric intake, has created a marketplace for more nutrient-dense, calorie-reduced foods. To meet this consuming need, the food industry must give nutritional criteria a prominent role in the product development process and optimize food products to meet the nutritional needs of Americans while keeping calorie levels low enough to match our less active lifestyles. Naturally, we are not willing to give up sensory qualities, low cost, and convenience, and the problem suddenly becomes so complex that a computer is needed to keep track of the multiple criteria. Using a comprehensive nutrient databank and a variety of software, we can now have the computer do the tedious work of determining the combinations and levels of ingredients needed to maximize nutritional value while minimizing calories and costs.

To be most efficient, the system must integrate the databases containing the various types of information needed, to enable simultaneous consideration of the actual eating habits, health status and food preferences of the consumers as well as the sensory, cost and nutritional characteristics of the potential ingredients.

This poses a challenge both to the expertise of the software creators and to the imagination and daring of the food technologists and nutritionists as well. The original role of the nutritionist in diet evaluation and menu planning is now being expanded by the need to understand computer languages and computer systems.

"THE NUTRIENT DATABASE: LABELLING AND PRODUCT DEVELOPMENT*

N. Rawson and C. S. Khoo
Campbell Institute for Research and Technology

Nutrient databases have become more important in the food industry since the creation of the nutrient labelling regulations in 1973. Whereas the original purpose for nutritional labelling was to help prevent deficiencies and promote awareness of ingredients, current concerns center on substances to avoid such as fat, cholesterol, sugar, and salt. Thus, industry has the government, health professionals, and the consumers to consider, and America's interest in and pursuit of "good nutrition" has never been greater. Today the low-calorie, nutrient-dense foods represent one of the fastest growing markets in the country. This has put tremendous pressure on the food industry to optimize its product development process through the use of much more comprehensive and innovative nutrient database systems.

My presentation today will focus on the general characteristics of the databases used in a food company, as well as how they function as a vital tool in many areas of the company, particularly in the labelling, product research and development functions.

There are two main types of nutrient databases commonly used by the food industry; an internal database, covering company products, and an "external" nutrient database, containing data on other food items (Fig. 1). The internal monitoring database contains nutrient data on the company's products based on laboratory analyses. The label values are derived from the analytical data and stored in a labelling database.

Federal regulations state that complete nutritional labelling is triggered by the addition of nutrients during processing or by any specific nutrition-oriented claims on the label or in advertising. When labelling is required, values for calories and ten nutrients: protein, fat and carbohydrate; vitamins A, C, thiamin, riboflavin, niacin; iron and calcium must be listed using the units shown, and as a percent of the U.S. Recommended Daily Allowance. Percentages stated must be based on the U.S. Recommended Daily Allowance, and expressed on a per serving basis (Table 1). Values for approximately 15 other nutrients (vitamins B6 and B12, phosphorous, magnesium, zinc, copper, biotin, and pantothenic acid, cholesterol, fiber, sugar) are optional.

Under current regulations, an optional nutrient, such as phosphorous, can be listed without triggering full nutritional labelling. However, if any required nutrient is listed, they all must be.

As research continues to point out the importance of specific nutrients in relation to the health concerns of today's population, the demand for increased labelling grows. I'm sure most of you are aware of the proposal currently before the FDA requesting mandatory sodium labelling, and I'm sure others will follow. Thus, while the minimum number of nutrients of concern in the labelling process is anywhere from 10 to 25, this number is increasing constantly.

At Campbell's we currently have data for about 1,000 products, and this data must be monitored several times a year, by performing laboratory analyses on representative samples from each plant. This amounts to at least 30,000 individual analyses done annually in our lab, and a large and constantly changing database of analytical values.

The monitoring database contains all the information needed to calculate the label values (Fig. 2). The procedures used are defined by the FDA in the Code of Federal Regulations (or CFR). The computer performs calculations of percent U.S. RDA and estimated Protein Efficiency Ratio, or PER, which is needed to calculate the percent RDA of protein. Statistical tests are done to set the label value.

The label values are then stored in a separate file and updated periodically.

While these regulations help to simplify and standardize the nutritional information presented on a label, it is the continually updated database which will inevitably be a company's primary resource for product data for use in education, advertising, and in product research and development (Fig. 3).

As accurate and complete as this dynamic database is, however, another source of food composition data, which I've called the external database, is needed for several reasons.

When a product reformulation is being considered, a source of nutrient data is needed to estimate the nutritional composition of the new product without the time and expense of laboratory analyses. This calculated value can then be compared with the current label, to see if label changes would be necessary. It would also be compared with nutritional standards to ensure maintenance or improvement of the nutritional quality of the product. This information would then be incorporated into a cost/benefit analysis to help plan the reformulation project.

A similar process is followed in the case of new products, and I'll be discussing this in more detail in a few minutes.

An external source of data is also useful in assessing how well a product compares with similar products available to the consumer--either homemade, restaurant or manufactured products. This can be used both as a guide in product development or in the marketing of existing products.

Thus, a source of external food composition data is needed, but, as this database is generally only updated once a year, to be most effective it must be linked to the internal database which will always contain more accurate data on company products. In order to do this, the database system must be flexible enough to enable this merging to occur as well as allowing easy access for frequent updating.

The "external" database must meet a number of criteria in order to be most useful to a food company. Use of the database in the research and development of new products and in advertising and product promotions imposes certain requirements different from those which may be needed for other purposes such as diet evaluation. I would now like to describe these requirements and explain why each is important in the research and development process.

The food items should include not only ingredients found in the kitchen such as raw egg white, but also commercial product ingredients such as "powdered egg white" (Table 2). Currently, most data banks have been developed for diet evaluation or menu planning, and commercial ingredients have largely been ignored.

In addition to raw ingredients, foods packaged and prepared in a variety of ways should be included, to enable imitating the effects of processing as much as possible.

The database should also include brand name items, recipes from prominent cookbooks, and restaurant items in order to insure that the nutritional quality of the current or proposed product at least equals that of similar products. Some companies are particularly interested in certain ethnic or international markets, and the database should be assessed with that in mind as well.

The nutrients listed in the database should include at least the 25 listed previously in relation to food labelling and for which there are RDA's set by the National Academy of Sciences (Table 3). Also, nutrients which may be of concern to specific target populations should be included. These might include other minerals and electrolytes such as flouride, selenium, and potassium. Amino acid data is also useful to enable estimation of the protein quality of a proposed product.

There are a number of other food components which may also be of interest in product development. These include components which certain groups may need to avoid such as allergens like lactose and gluten, as well as food characteristics such as digestibility or protein efficiency ratio. Where accurate quantitative data is difficult or impossible to obtain, a flagging system would be very useful for keeping track of this type of information during the product development process.

Another aspect of the database critical for any purpose but particularly important in product development, is the amount of missing and assumed data. When designing a product based on nutritional criteria, the relevant nutrient data must be complete.

It is important that missing data be identified as different from assumed or true zero values.

Nutrient data should be available for both 100g units or typical serving size, and the values used for converting from a volume quantity such as a cup, or a typical serving to equivalent weights must be listed and documented. We have found that this conversion is a common source of error In nutrient composition calculations, partly due to imprecise definitions of an average serving, for instance, and partly due to assumptions made in the factors used to convert between volume and weight.

In addition, the source of each nutrient value should be identified. The reference should be traceable and include the method of analysis. This is particularly important when developing promotional and advertising material as it is subject to the scrutiny of not only the corporate scientists and lawyers, but the legal departments of the advertising agency and the media as well.

Another aspect important in product development is the coding system used. The food codes should enable accessing the data by food group such as "meats" as well as food type such as "beef". This is useful in the early phase of product development where one may wish to compare potential ingredients based on nutrient composition. It would also be useful to have a code indicating the type of packaging and the state and method of preparation. A typical request during the product development process might be to provide a list of cooked roast beef items with more than 10% of the RDA of selected nutrients, per one ounce serving. With an efficient coding system and the appropriate software, this task is very easy to perform.

The software needed for maintenance and utilization of the database is often available for purchase. Because of the potential problems with computer systems compatibility, however, we have found it more efficient to develop much of our own software designed specifically for our own needs. This approach requires a programmer available on a continuous basis, at least while the system is being developed. The maintenance software must create an interface between the nutrient database and the internal monitoring database to insure that the most accurate data possible is used. One may also want to consider interfacing the nutrient database with other types of food data used in product development, such as the ingredient cost file. Be aware, however, that these other datafiles will have been established years ago with rigid designs geared for a specific purpose. Extensive support may be required from the programmers and computer systems analysts to develop these interfaces.

The applications software needed for product development covers a wide variety of areas, from simple nutritional calculations to complex statistical analyses, and involves virtually every area of the company. Some types of applications software can be purchased or rented, such as SAS or SPSS, and some will have to be developed for your particular needs. Computer science expertise will be needed in either case, however, to link the software with the databases and assist in the development of an efficient system.

Once developed, however, the benefits will be felt by the entire company in terms of efficiency and innovation.

The traditional product development process began with a concept initiated, most often, from the marketing unit (Fig. 4). The concept was then refined through discussions between the marketing personnel and the food technologists. The refined concept would be tested at consumer focus groups to determine the level of consumer interest. Prototypes would then be produced in the test kitchens and evaluated for sensory, cost, and engineering considerations. Revisions would be suggested, and the process would be repeated as many times as necessary. When the product was good enough to satisfy the internal inspections, it would be sent to test markets to determine consumer acceptance. When the product was finally ready for large-scale marketing, the product would be fully analyzed for nutritional labelling. Only then would the nutritionists be consulted to determine what nutritional strengths the product might have for use in advertising, promotions, and other consumer communications.

Two main factors have made this traditional approach inappropriate. First is the extensive progress made in the last decade in delineating the complex relationships between diet and health. Second is the high level of consumer awareness of the nutritional Quality of their diets. Thus, the current goal at Campbell's is to include nutritional evaluation right at the beginning of the product development process. In addition, nutritional criteria are set for new products based on the nutritional requirements of the proposed product users. The nutrient database has become a powerful and indispensable tool in the implementation of these goals.

I would now like to show you exactly how this is done, beginning with the first phase, or the development of a new product concept. While this concept can originate from several sources, the nutrient database can be used here in conjunction with the traditional methods to help identify and focus in on new product opportunities (Fig. 5).

The nutrient database enables a company to analyze diet surveys performed by the company or by others. The critical questions asked constantly by the market researchers, such as:

- "What kind of people eat or do not eat product A?" can be augmented by

- "How nutritionally healthy are their diets?"

- "Where does product A fit in?"

Say, for instance, it is found that the product user's diet was low in calcium, and that product A was not contributing a significant amount of that nutrient. Reformulations could be undertaken to increase the amount of calcium in product A, or to develop a complementary or substitute product which would provide the calcium needed. Naturally, there are many other factors that must be considered as well before such a project is undertaken. Other types of information can be obtained from the survey, such as demographic, anthropometric, biochemical, food preference and meal pattern data. All of this information can be used to help develop a product which fits the lifestyle and nutritional health characteristics of the proposed product user. Conversely, the database can be used to evaluate the product concept in terms of its potential impact on the user's diet and nutritional health. This is an important consideration in deciding which concepts warrant further development and which do not.

This nutritional evaluation can also be used to refine a concept into a high quality, saleable product. Nutritional criteria are set up, based on the U.S. Dietary Guidelines and the NAS-NRC recommended daily allowances for the target population (Table 4). In this example, based on typical guidelines for a frozen dinner, levels were set for calories, sodium, and specific nutrients as shown. Usually from this point many prototypes are developed, based on food preferences and technical considerations. One example is shown here and contains chicken, chicken broth, pineapple, green beans, rice and a cookie (Table 5). This information would be fed into the computer to calculate the nutritional composition of the meal. We then check to see how well it compares with the criteria stated previously concerning the levels of calories and nutrients.

We simultaneously calculate the percent U.S. RDA's, and check these against our criteria as well. Here, we achieved the 30% goal for only 4 nutrients, and had more calories from fat and protein than was desired. We also monitor the RDA values for the other age groups, to see that the product could be used by anyone without adverse effects. All this can be done in a matter of minutes, compared to the long hours of manual calculations or laboratory analyses required otherwise.

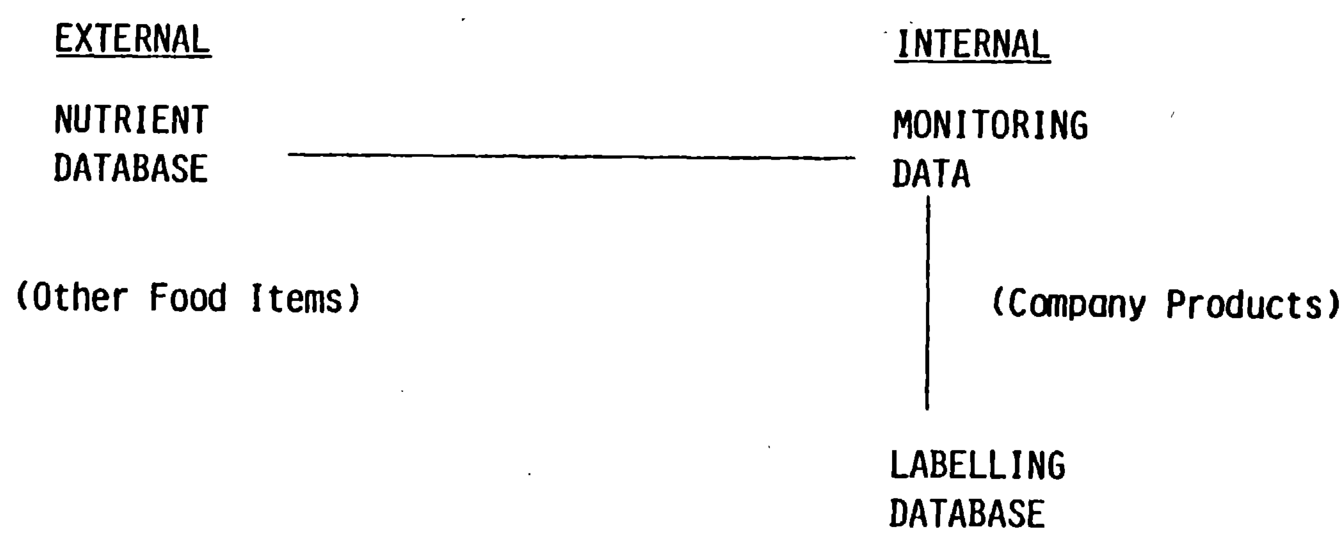
With the appropriate software, the database also helps us to reformulate the product by providing alternative ingredients or proportions that will help meet the criteria specified (Fig. 6). The new product continues through the sequence of inspections and reformulations as before, with the addition of a nutritional evaluation at each step, to be sure that the nutritional standards are maintained.

Only when the product is ready for the test market would it be analyzed by the laboratory.

Another type of product development research involving the nutrient database is in the area of enrichment and fortification. Whereas specific nutritional deficiencies are relatively rare in this country, concern has recently been expressed regarding iron and calcium status, particularly among women. The impact of fortifying or enriching a product can be assessed by analyzing internal or external diet survey data with and without the fortified product. The database allows us to see the effect on the user's total diet and predict positive or negative interactions. In view of the current widespread use of supplements, and the concern about nutrient excesses, this type of evaluation has become even more important.

Overall, what we have a picture of is a multidimensional database, integrating many types of data for a variety of purposes (Fig. 7). The research and development applications pose some specific problems and impose specific criteria on the system. I hope I have given you some insight into both the benefits and potential problems associated with the use of a nutrient database for enhancing the product development process in a food company.

FIGURE 1.



EXTERNAL
NUTRIENT DATABASE
(Other Food Items)
INTERNAL
MONITORING DATA
(Company Products)
LABELLING DATABASE

TABLE 1.

LABELLED NUTRIENTS

	Required	Units/Serving
1.	Calories	Kcal
2.	Protein	g, % USRDA
3.	Fat	g
4.	Carbohydrate	g
5.	Vitamin A	IU, % USRDA
6.	Vitamin C	Mg, % USRDA
7.	Thiamine	Mg, % USRDA
8.	Riboflavin	Mg, % USRDA
9.	Niacin	Mg, % USRDA
10.	Calcium	g, % USRDA
11.	Iron	Mg, % USRDA

FIGURE 2.

INTERNAL DATABASE

Monitoring Data

- Product No.
- Laboratory Means
- Standard Error
- Sample Size
- Plant No.

--Continuously Updated

Software

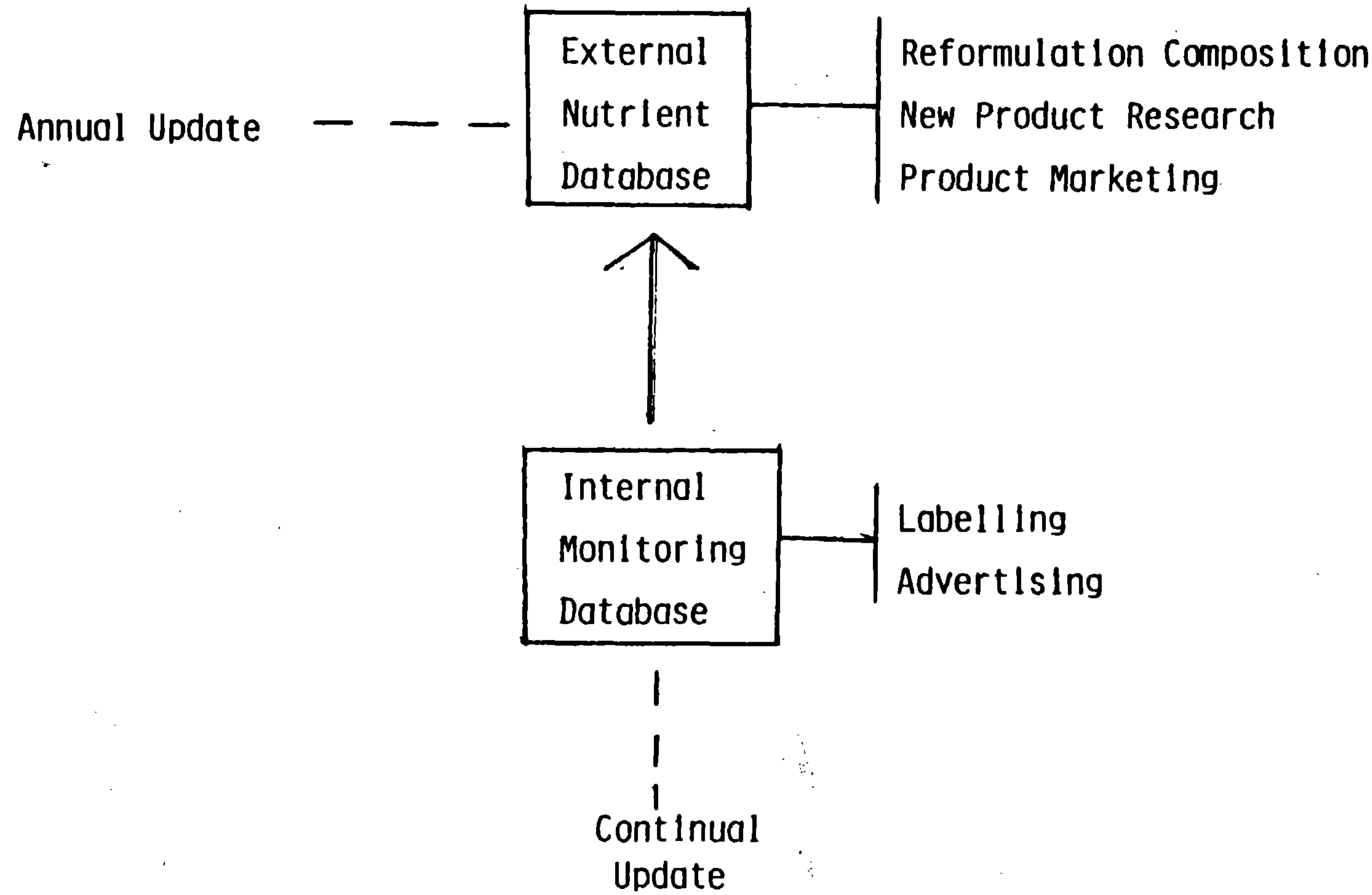
- CFR Regulations
- % USRDA
- PER
- Confidence Level

Lins

- Product No.
- Current Label
- Proposed Label
- Risk Factor

--Updated Quarterly/Annually

FIGURE 3.



Annual Update
External
Nutrient
Database
Reformulation Composition
New Product Research
Product Marketing
Internal
Monitoring
Database
Labelling
Advertising
Continual
Update

TABLE 2.

EXTERNAL DATABASE

FOOD ITEMS

1. Raw Ingredients
   - Retail, i.e. raw egg white
   - Wholesale (Commercial)
     i.e. powdered egg white, stabilized
2. Cooked Foods

- Boiled	- Sliced
- Broiled	- Mashed
- Steamed	- Whole

3. Prepared Foods
   - Brand Names
   - Recipes
   - Restaurant Items

TABLE 3.

EXTERNAL DATABASE

NUTRIENTS

1. Labelled Nutrients (10-25)

2. Other Nutrients
   - Trace Minerals
   - Electrolytes
   - Amino Acids

OTHER COMPONENTS AND CHARACTERISTICS

1. Caffeine
2. Allergens - Lactose, Gluten
3. Digestibility, PER

FIGURE 4.

TRADITIONAL PRODUCT DEVELOPMENT PROTOCOL

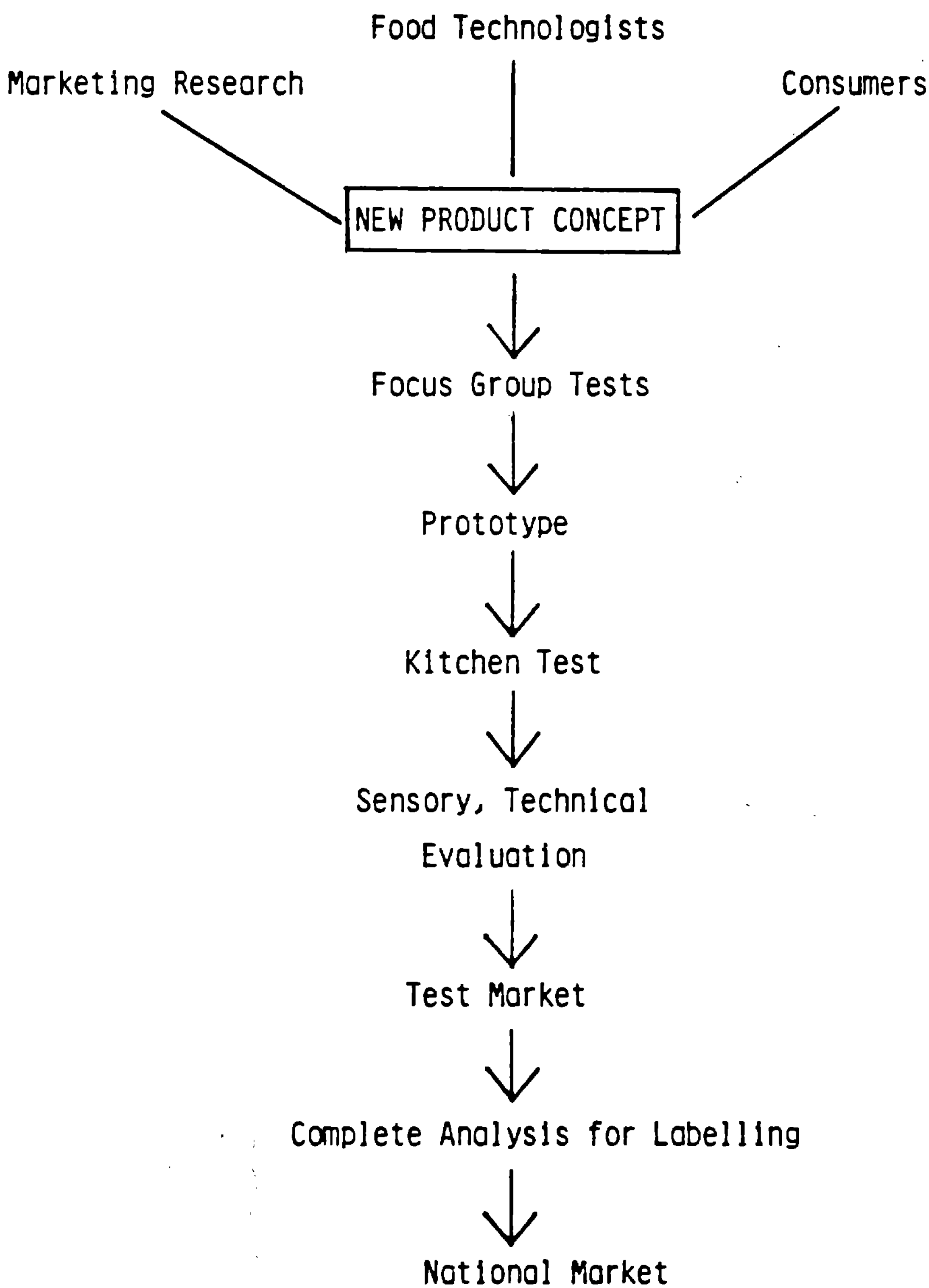


FIGURE 5.

USE OF NUTRIENT DATABASE IN PRODUCT DEVELOPMENT

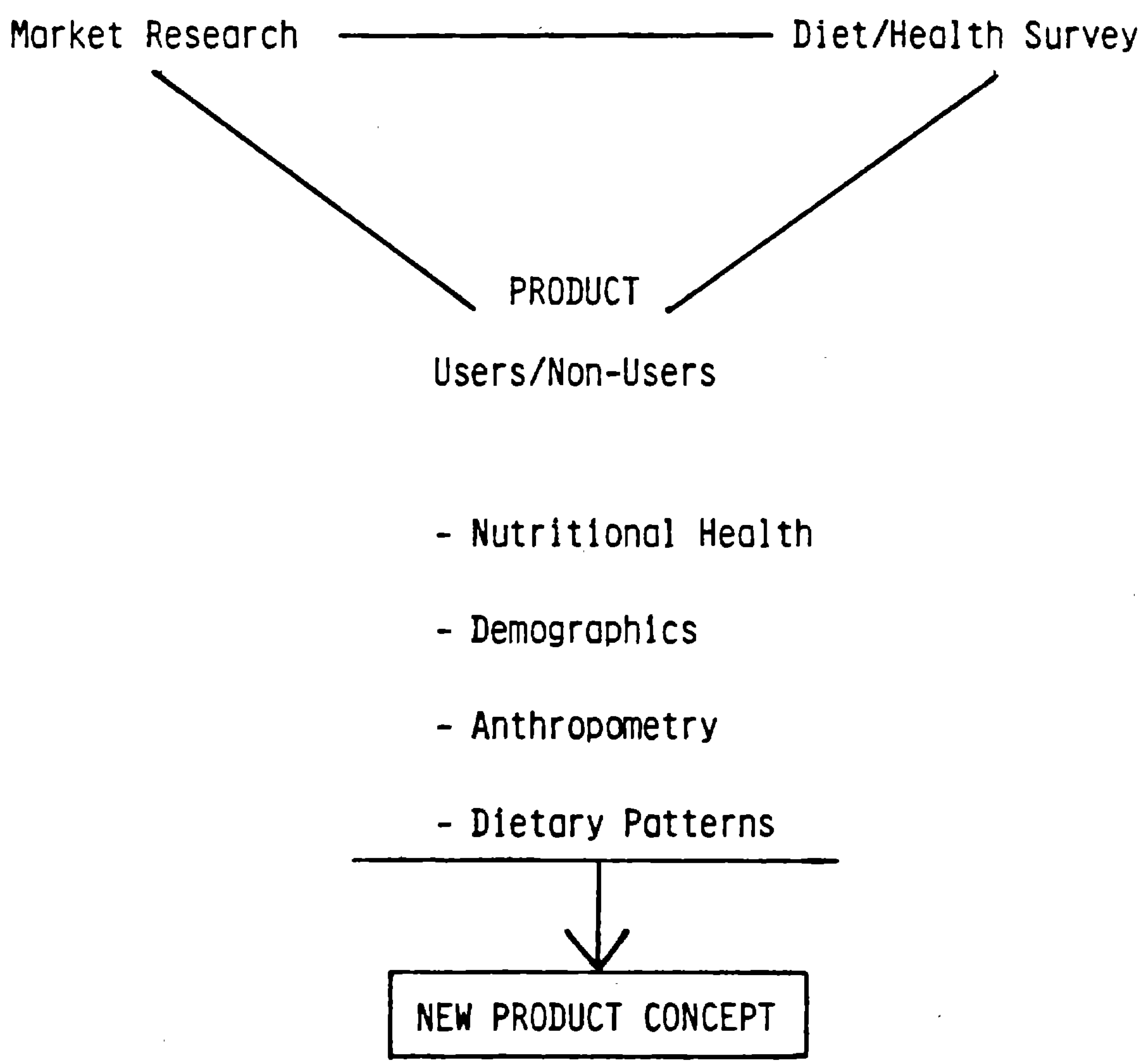


TABLE 4.

EXAMPLE

NUTRITIONAL GUIDELINES

1.	Calories	1/4 - 1/3 of the RDA (Adults) - 500 - 700 Kcal
2.	Sodium	Less than 1000 mg
3.	Nutrients	At Least 30% RDA
4.	Fat	Not More Than 30% of Calories
5.	Protein	12 - 15% of Calories
6.	Carbohydrate	50 - 55% of Calories

TABLE 5.

SAMPLE

SAMPLE	WT. GRAMS
CHICKEN, BROILER/FRYER, BREAST, MEAT ONLY	94.0
SOUP, DRIED, CHICKEN BROTH, CUBED,	1.8
LEMON JUICE, FRESH	5.0
PINEAPPLE JUICE	31.0
PINEAPPLE CUBED	65.0
GREEN BEANS	56.8
MARGARINE, LOW SODIUM	4.9
BROWN RICE CKD	14.2
FORTUNE COOKIES	0.5
SUGAR, BROWN	0.6
VINEGAR, CIDER	10.0
ALMONDS, CHOPPED	1.8

FIGURE 6.

USE OF NUTRIENT DATABASE IN PRODUCT DEVELOPMENT

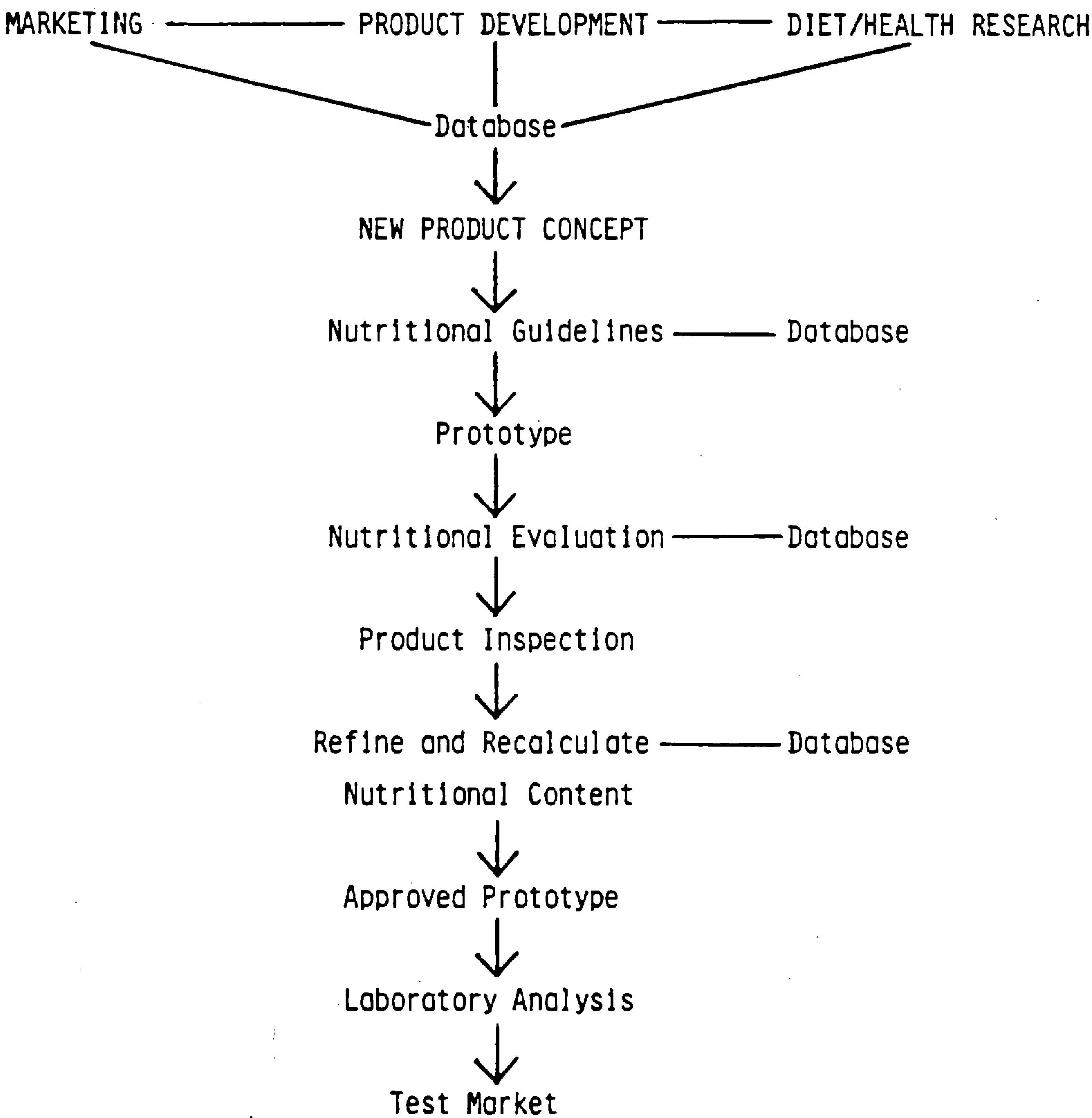
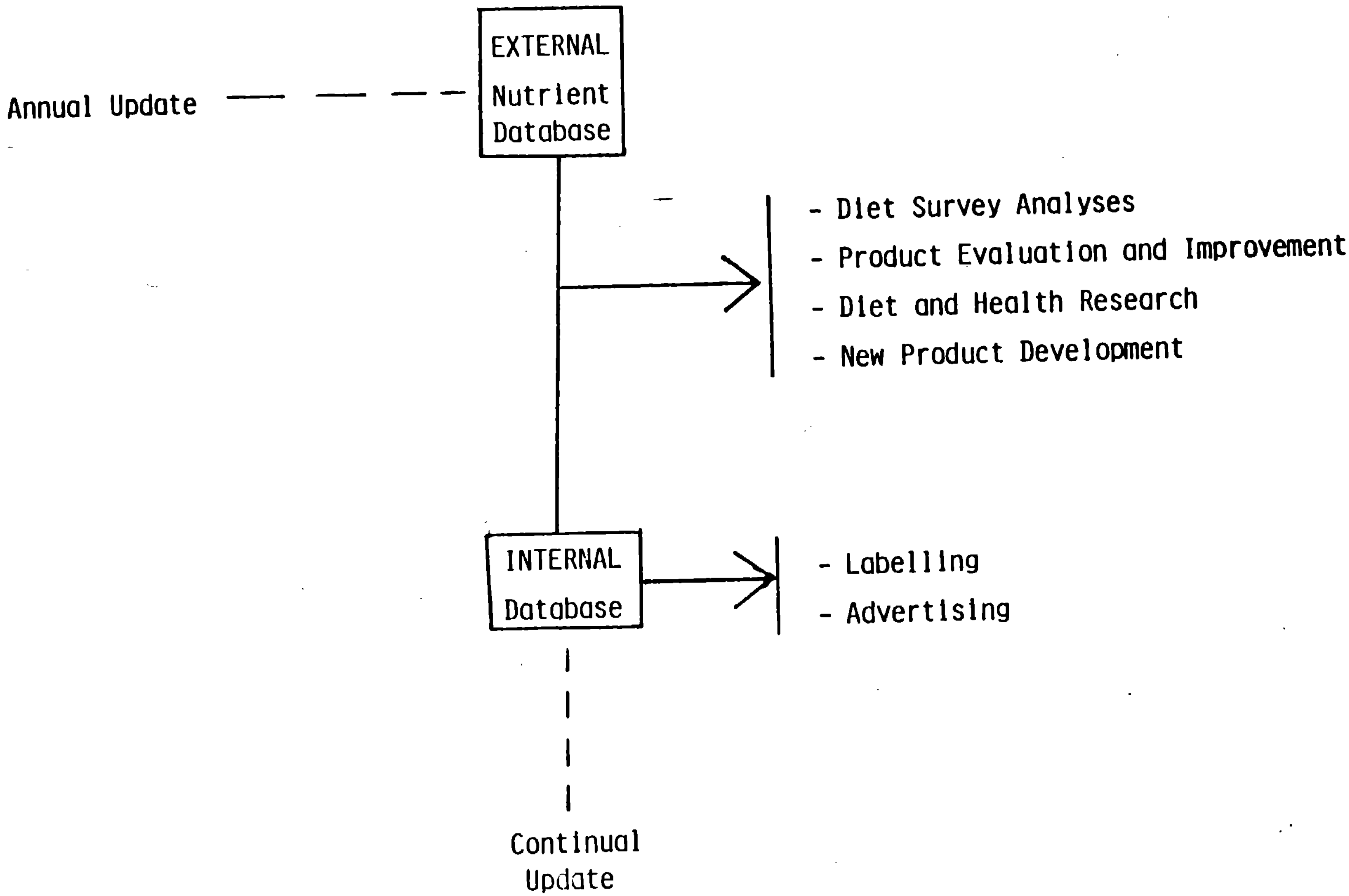


FIGURE 7.


Annual Update
EXTERNAL
Nutrient
Database
- Diet Survey Analyses
- Product Evaluation and Improvement
- Diet and Health Research
- New Product Development
INTERNAL
Database
- Labelling
- Advertising
Continual
Update

LARGE-SCALE DIETARY ANALYSIS AND EVALUATION

Bill R. Sanford, President
Nutrition Services Division
Health Development Inc.
1165 West Third Avenue
Columbus, Ohio 43212

INTRODUCTION

Thank you for the introduction, Dr. Samonds. It is a real pleasure to be able to speak today at this pre-conference program on the usage of a nutrient data bank for large-scale dietary analysis and evaluation.

Two excellent articles on the use of data banks are in the May, 1983, Journal of the American Dietetic Association by Dr. Loretta Hoover and Betty Perloff.

When Dr. Samonds asked me to make this presentation, he indicated that my comments should be geared toward the novice in data bank usage. This certainly makes me feel more comfortable since my presentation will not be highly technical. Rather, I will talk about Nutrition Services' commercial application of a computerized diet analysis program using a well documented nutrient data bank.

The best way to talk about our usage of a data bank for large-scale analysis and evaluation is to describe our business. I will try to be non-commercial, but it is difficult because I am so proud of our company and what we do.

Health Development Inc., the parent company of Nutrition Services, is involved in two of the most exciting fields in American business--computer products/services and health services. As pointed out in John Naisbitt's book Megatrends, our economy is changing from an industrial base to an information base. We at Health Development are working to assist health professionals to more effectively and economically utilize health-related information, including nutrient analysis data.

Our company became interested in the possibility of a nutritional assessment/dietary planning computer program in the latter part of 1979. We saw a large potential market for nutrition information. I could talk for hours about the reasons for the large and growing market, but today we will stay on the subject of large-scale dietary analysis and evaluation. However, you might be interested in what we see as some of the potential users of services and/or computer programs related to nutritional assessment and dietary planning.

Potential Users:

250,000	M.D. Physicians in private practice
40,000	Registered Dietitians
7,000	Hospitals
21,000	Nursing Homes
17,000	D.O. Physicians in private practice
18,000	Chiropractors
150,000	Dentists
5,000	Weight Loss Centers
500	Corporate Fitness Centers

When we began our search for some sort of nutrition analysis "thing," we were not sure exactly what we wanted. However, one thing was quite apparent at the time. There were only two basic ways of doing diet analysis calculations--manually or with a very large computer system.

Our search for a program led us to Dr. Sarah Short. Most of you know of Dr. Short because of her innovative teaching techniques at Syracuse University where she is Professor of Nutrition. She is also nationally recognized for her many years of work with computerized nutrition analysis programs for her research and student education. Some of her most interesting research work has been on nutrition and athletes.

I first became familiar with Dr. Short's work while watching the TODAY show on television. She had a computer terminal with her and was doing an analysis of Tom Brokaw's diet. The application was intriguing, so I contacted Dr. Short a few days later to discuss some commercial opportunities.

The SHORT REPORT® Computerized Diet and Energy Analysis was introduced in March, 1980. The SHORT REPORT® is a complete service package designed for use by health professionals in the evaluation of patient/client food consumption and physical activities. The applications of the analysis program range from those related to diagnosis/treatment to health promotion/wellness/ fitness applications. The analysis is most commonly used by health professionals who are working with non-hospitalized individuals. Our company currently has users of the service in 36 states and Canada. Thousands of nutrient analyses have been performed by our company, so I guess that qualifies me to be talking about today's subject.

When our company first became interested in computerized nutritional analysis, we set some criteria for the food and nutrient data bank that would be used. The criteria were fairly simple in concept, but have remained important to us. The following are the criteria:

1. The food and nutrient data bank must be well documented.
2. Must contain a large number of foods and serving sizes.
3. Should have the fewest nutrient "holes" as possible (for example, we do not feel that a data bank that includes trace minerals such as selenium has much application for us because of the small number of foods that have been analyzed for selenium).
4. Must be routinely updated.

Fortunately, Dr. Short's data bank met all of these criteria. The primary source of food and nutrient information in the data bank we use is the U.S.D.A. This information is augmented with published, well documented research, and numerous brand names and fast foods. Our data bank presently has over 5,000 foods and serving sizes entries.

While the data bank is extremely important, there are two other equally critical considerations in having an effective nutrient analysis package. A well written and well documented computer program is a must. The analysis process must also have an effective way of collecting food and beverage consultation information. The SHORT REPORT® Computerized Diet and Energy Analysis includes all three of these important aspects--sound data bank, well written programs, and effective data collection.

NUTRIENT ANALYSIS PROCESS

The complete nutrient analysis process utilizing The SHORT REPORT ®begins with the health professional. This individual is most likely involved with patient/client evaluation and counselling. We prefer to deal with health professionals because they have a better understanding of both the analysis process and the final data developed through the analysis.

The individual whose diet is being analyzed is provided with a Patient Data Kit containing all necessary instructions and forms for recording dietary intake and, if appropriate, physical activities. Our standard package is a three day analysis, but we can handle up to a 14 day period.

The recording logs are designed for ease of use by the individual. Each log contains a listing of 370 of the more commonly eaten foods and common serving sizes. The log is designed to be carried with the individual so that food consumption can be recorded as it occurs. When someone consumes a food that is not preprinted on the list, the food description and serving size are written in a special "additional foods" space on each daily log.

When the recording is completed, the daily logs are mailed by the patient (or sometimes the health professional) to Nutrition Services in a prepaid, self addressed envelope that is included in the Patient Data Kit. Each log is reviewed and coded for computer processing under the direction of a Registered Dietitian.

Patient demographic data, food consumption data, and physical activity information is entered into the computer and a multipage report is printed which includes the following information:

1. A listing of the foods, serving sizes, and numbers of servings recorded by the patient.
2. A detailed breakdown of food/beverage nutrient information for each day of the analysis period, including suggested intake levels, actual intake, percentage comparison of actual to suggested, and intake from supplement sources.
3. A nutrient summary with a graphic representation for the entire reporting period.
4. A detailed energy summary for each day of the analysis period, including total energy requirement, actual intake, and the excess or deficit of actual to the suggested levels.
5. The text of summary notes that relate to specific excesses or deficits of actual nutrient intake to suggested values. The text includes dietary suggestions, sources of nutrients and energy, and potential effects of continued excesses or deficiencies.

Two copies of the completed report are returned to the health professional in a confidential envelope. As stated earlier, we prefer to deal with health professionals because they best understand the information included in the final report and the limitations of a nutritional analysis itself.

One important feature of our program that I didn't previously mention is the ability to handle non-data base foods in any individual analysis. A person will sometimes consume a food that is not in our large data base, but we will have information on the nutrient content of that particular food. For example, we may have a package label. The nutrient information for this food can be entered into the computer along with the data base foods to result in a more accurate total analysis.

I mentioned earlier that our company has performed thousands of analyses in the manner just described. However, the entire process is time consuming primarily because of the delays created by the mail service. The patient or health professional must mail forms to our office, we must process the information, and then return completed reports to the health professional. Our in-house processing time is one day, but the entire turnaround time can often be as much as seven days because of the mail.

Because of the turnaround time delays, we started looking at alternatives to our standard service package about three years ago. The most appropriate alternatives were some sort of time sharing system and the use of microcomputers. We did not really like the time sharing idea because of the perception of potential users of the difficulties of working with a large remote computer site. The most attractive idea was to develop a comprehensive analysis program for use on popular microcomputers. This is what we decided to do.

NUTRIENT ANALYSIS EVOLUTION

We began a project to convert our program and data base from a large computer to a microcomputer. At that time we had versions of our programs running on both an IBM 370 and a DEC PDP-1170. The biggest problem in the conversion process was getting our large data base on a single floppy diskette that could be used with something like the Apple II microcomputer. We felt strongly then and feel equally strongly now, that a data base of only a few hundred foods was not sufficient to provide a meaningful analysis.

A two year development program resulted in The SHORT REPORT ® MICRO, a microcomputer software program for diet, menu, and energy (caloric) analysis. The new program is truly remarkable in that it has actually more analysis options than we had on the big computers. The following are some of the features of the microcomputer version:

1. Eight analysis options including a food analysis which shows the nutrient composition of each food item.
2. Over five thousand foods and serving sizes in the data base on a single floppy diskette.
3. Analysis of thirty-three nutrient and energy components.
4. The ability to add up to 700 additional user defined foods to the permanent data base on the IBM PC version of the program.
5. Versions for the Apple II+, Apple IIe, Apple III, and IBM PC (a new IBM PC/XT version has just been released).

The health professional can now do a comprehensive nutrient analysis in his/her own office. The program is fast, easy to use, economical, and accurate. The most positive feature is the rapid availability of detailed nutritional analysis information for use by the health professional.

SUMMARY

I have given you a brief overview today of an example of the use of a nutrient data bank for large-scale dietary analysis and evaluation. There is a large and growing market for detailed nutritional information if such information can be provided in an efficient, accurate, economical, and rapid manner.

Nutrient analysis has evolved from manual systems, to large computer systems, and now to microcomputers. Regardless of the method of analysis, the choice of a data bank is an important consideration. Other considerations should be the method of food/beverage consumption data collection and the software programs themselves when a computer is being used.

Thank you very much for allowing me to speak with you today. Our microcomputer version of the analysis program, The SHORT REPORT ® MICRO, can be seen in the exhibit area through tomorrow. If you have any questions about our nutrient analysis service or software programs, please feel free to call us at 1-800-222-4630.

Thank you again.

STATUS REPORT ON USDA NUTRIENT DATA

F. N. Hepburn
Chief, Nutrient Data Research Branch
Consumer Nutrition Division
U.S. Department of Agriculture

The stated mission of the Nutrient Data Research Branch is to compile and make available data on the nutrient composition of foods. This includes developing and maintaining the Nutrient Data Bank and publishing tables of food composition. Our operational objectives include making available up-to-date information on the nutrient composition of foods by publishing revised sections of Agriculture Handbook No. 8, encouraging the generation of new, reliable food composition data, and cooperating with nutrient data users and suppliers. A most important facet of cooperation is communication and we welcome the opportunity afforded by this annual meeting to exchange information and discuss the mutual problems we face in dealing with nutrient composition data.

The purpose of this presentation is to describe the progress being made on the revision of Agriculture Handbook No. 8 and the availability of provisional data and to discuss the present state of knowledge of food composition and efforts being made to fill gaps in that knowledge. I shall also discuss briefly some of the problems faced in trying to obtain reliable information on specific food components of current interest.

REVISED SECTIONS OF AH-8

As most of you know, nine sections of Agriculture Handbook No. 8 have been published to date. These are listed in Table 1. Numbers 1 through 7 are still available from our office without charge to professional workers but because of the new publication policy instituted last year, copies of all succeeding sections are available only for purchase from the Government Printing Office. These nine sections have been incorporated into the USDA Nutrient Data Base for Standard Reference, Release 3, 1983, (as described in the handout), Datasets Available in Machine Readable Form. On the computerized tape, imputed values have been supplied, where possible, to fill in the missing values in the printed sections.

Work is proceeding on all remaining sections and their present status is indicated in Table 2. Section 10, Pork Products, has been sent to the printer and should be published in a few weeks. This section incorporates data from an extensive study conducted by the Meat Science Research Laboratory and Nutrient Composition Laboratory in cooperation with industry and reflects the composition of pork as marketed today. Section 11, Vegetables and Vegetable Products, is in review. This section was greatly expanded over the 1963 edition of the handbook and will include nearly 500 items. Publication is expected this winter. Numbers have not yet been assigned to the remaining sections and the order of publication may differ from the sequence indicated in the table. We do, however, expect the first four or five of these to be completed in 1984.

In Table 2, existing data are at the Data Base II level for Legumes and Baked Products with a note indicating that additional data are expected to be received before the sections will be completed. Actually, additional data are expected for food items in each of these food groups as a result of research efforts now being conducted under contracts, grants, and cooperative agreements through our agency, by the Nutrient Composition Laboratory and by the Meat Science Research Laboratory. These efforts are directed towards filling gaps in our knowledge and are explained below in the discussion of Current State of Knowledge.

PROVISIONAL TABLES

To help meet the demand for up-to-date information, we have been releasing provisional tables of data for specific nutrients or for limited numbers of foods for which sections have not yet been revised. These data are to be considered strictly provisional and will be superseded by the information published in the completed sections. Tables available since our last meeting are shown in Table 3.

The Iron Content of Foods may be purchased from the Government Printing Office, but copies from our limited stock will be mailed to those who send written requests, as long as our supply lasts. Last year we described a unique feature of this table, which was the inclusion of confidence codes along with the data, in an attempt to describe the degree of reliability attached to the values. We shall appreciate receiving your comments and suggestions about these codes. Do you find them helpful? Can you suggest ways to make them more useful? The second two tables listed are printed on single sheets and are obtainable upon request from our office only.

In preparation, and scheduled for release by the end of this year, is a Provisional Table on the Carbohydrate Fractions of Foods. This will be a compilation of existing data on sugars, starch, and complex polysaccarides.

CURRENT STATE OF KNOWLEDGE

Table 4 depicts the current state of knowledge of the nutrient composition of foods. This information was updated in January of 1983 by the staff of the Nutrient Data Research Branch. This type of presentation has been exhibited in previous years and it was most gratifying to see the changes that could be made this year in upgrading the status of many nutrients in many food groups. An innovation in Table 4 is the inclusion of indicators to show areas where research is underway to generate needed data.

For the most part, these areas represent ongoing and planned work sponsored by the Consumer Nutrition Division in the form of contracts, grants, and cooperative agreements with land-grant universities and non-profit research organizations. They also include some of the work performed by the Nutrient Composition Laboratory and the Meat Science Research Laboratory at Beltsville.

I do not want to give the impression that all of our problems will soon be solved as a result of these activities. We know that these efforts will affect genuine progress towards filling in gaps in our knowledge of nutrient composition, but the studies are necessarily limited in scope and still will leave many boxes coded as "insufficient" information. For this reason, I shall give a brief description of some of the studies to outline their scope and limitations.

Most of our efforts are food oriented. That is, we are seeking information for particular foods on all nutrients listed in the revised sections. Our efforts are concentrated on food groups for which sections are still in preparation. Some studies, however, are nutrient oriented because of the need to develop data requiring special methodology across many types of foods. These are discussed separately.

Studies on Foods

Baked products and cereal grains: Eight variety breads, tortillas, pita bread, bagels, and three pastas, raw and cooked, are being sampled from eight cities. Two-hundred products, including breads, cakes, pastry and hot cereals, are included in a separate study.

Beverages: Beverages are being studied as components of a group of 50 miscellaneous foods.

Fish: Six raw fish and eight processed fish are under study. A separate project will determine the total lipid and cholesterol contents of 10 mollusks sampled at different times and locations over 2 seasons.

Fruits: Six fruits sampled from six cities are being analyzed.

Legumes: An extensive study on legumes is nearing completion. Eighteen legumes are being analyzed in raw and cooked forms, sampled from four growing locations, before and after 1-year's storage. Another project involves the nutrient content of unsprouted and sprouted cereal and leguminous seeds. This covers 13 legumes from four growing locations over 2 crop years and six sprouted seeds, both commercial and laboratory germinated.

Meat: Major studies on pork and on beef have been completed. An examination of ground beef is underway. Studies on lamb and veal have just been completed. These projects do not include variety meats. This is the subject of a separate project in which nutrients are being determined in five variety meats (brain, heart, liver, kidney, and tongue) from three species (beef, lamb, and veal). We also need to know the consequences of cooking meat and poultry, after removal of fat. Research will be initiated soon to determine nutrients after cooking representative cuts of beef, pork, and poultry from which fat and/or skin had been trimmed prior to cooking (four cuts, three species, and four animals of each species).

Nuts and seeds: Some work with legumes and cereals has been mentioned. Other work is included in the miscellaneous foods studies.

Vegetables: Work includes 12 items from six geographical regions, analyzed raw and some after cooking. The distribution of nutrients between the drained solids and liquid portions of 96 samples of canned fruits and vegetables is also under investigation. The effect of particle size (whether whole, sliced, chopped, etc,) and of the canning medium is also studied.

Mixed dishes: Fifty mixed dishes, analyzed raw and cooked, are being studied to measure retention of nutrients during preparation and cooking. Results will test the reliability of estimating nutrients in recipe-formulated foods from the list of ingredients.

Fast foods: All nutrients, including amino acids and fatty acids, are being determined in 42 fast foods from available restaurants in a single location. Ten types of fish foods from fast service establishments and 10 frozen convenience fish foods are under study. Fifty-one Mexican-American fast foods, obtained from four fast-food restaurants are being analyzed. A study of the nutrient composition of fast-food fried chicken is getting underway in cooperation with the Nutrient Composition Laboratory (NCL). The NCL will collect approximately 100 samples on a nationwide basis and process them for analysis.

Home-prepared and institutionally-prepared foods: Six institutional foods from each of three food service establishments and 50 miscellaneous home-prepared foods are being analyzed.

Miscellaneous foods: Three separate contracts allow for the analysis of 150 different foods as prescribed. These foods include some from the truly miscellaneous category, but extend to food items in the beverages, candies, baked products, cereals, and nuts and seeds groups to help us fill specific gaps in the data.

Studies on Nutrients

The determination of cholesterol in mollusks was mentioned in connection with fish. Other studies planned to be started this year are on amino acids, vitamin A active compounds, and carbohydrate fractions. We are seeking to determine the amino acid composition of 500 food items, utilizing samples prepared for some of the above studies. This study is aimed primarily at the mixed dish and fast food categories in order to test the reliability of calculating amino acid contents from those of the ingredients.

Because of the current interest in beta-carotene, work will be initiated to determine alpha- and beta-carotenes, cryptoxanthin, and retinol in 25 representative foods of both plant and animal origin. Samples will be collected from five locations in three seasons of the year.

Carbohydrate fractions and fiber constituents will be measured in 112 samples, selected to take into account the effects of cooking and processing.

A final project that should be of interest is a review of the Atwater system for estimating the energy content of foods. This was undertaken by the Life Sciences Research Office of the Federation of American Societies for Experimental Biology under a grant from the Consumer Nutrition Division. A comprehensive review of the literature was completed and a workshop meeting of an ad hoc expert Committee was held to discuss the appropriateness and limitations of the present system. A final report, summarizing the Committee's deliberations and recommending procedures and changes for improving the system will be submitted later this year.

PROBLEM NUTRIENTS OF SPECIAL INTEREST

Dr. Gary Beecher, Chief of the Nutrient Composition Laboratory, has provided Table 5, updating the state of development of methods for nutrients in foods. In general, the nutrients included in the revision of Agriculture Handbook No. 8 are those for which the methods are considered to be adequate or substantial in terms of that table.

Because of the intense interest in knowing about certain nutrients and energy we have extended the collection of data for Handbook 8 to include (besides energy) some nutrients for which the methods may give conflicting results. These are manganese, vitamin A activity, vitamins B12, C, and E, folacin, and pantothenic acid. In our contract work we can specify methods and request that the laboratories utilize sound principles of laboratory quality control to validate the methods as well as possible. Chief problems reported to us concern folacin, for which most laboratories find a high variability among replicate determinations and for which we continue to observe a wide variation among laboratories on similar foods. We believe that our continuing meetings with collaborators working on our contracts are helping to solve some of the problems in food analysis but final resolution of analytical problems must be achieved before we can place highest confidence in the values reported for this list of nutrients, especially folacin. At the present time, we must report the data as being best available, recognizing that they are subject to change when improved methods are established and applied to representative samples of food.

There are severe problems with selenium and beta-carotene, two nutrients being linked to the relationship between diet and cancer, and there is an urgent need to develop new data. The method for selenium is considered to be substantial and although tedious, it is believed to yield reliable results. The problem is that the distribution of selenium is known to be so variable in the food supply that meaningful, representative values cannot be estimated from existing data. Selenium is known to vary with geographical region, animal feeding practice, and food distribution patterns. A major study is required to quantitate these variables and enable the development of a rational approach to the estimation of selenium in foods.

Until now, it has been common practice to report only the total vitamin A activity of foods, and not the separate contributions of retinol and pro-vitamin A isomers. Values for beta-carotene have not been reported frequently and existing reports are often not clear as to whether a value is explicit for beta-carotene or whether it may include other isomers. Furthermore, it is not clear whether the relationship to cancer is limited to beta-carotene, to other pro-vitamin A carotenoids, or to all carotenoids regardless of vitamin A activity. In other words, we are not sure that only beta-carotene data are truly needed for investigating the relationship of diet to cancer.

We believe that a major effort should be launched to determine the contents in foods of the separate carotene isomers. Toward this end, the Nutrient Composition Laboratory is in process of developing methods. The single project we are sponsoring on 25 foods should provide helpful information on the distribution of vitamin A and related compounds in several kinds of foods and mixtures. Until such definitive information becomes available, only a crude estimate of carotene can be made by applying partition coefficients to the total vitamin A activity. We are aware that some people are constructing data bases in this fashion, using the coefficients published by the Food and Agriculture Organization of the United Nations. Evidence in our hands, however, suggests that these coefficients may be incorrect and we are concerned that their application may result in misleading information.

We are frequently asked for data on dietary fiber. This continues to be a problem because fiber is defined quantitatively by the method used to measure it and, as yet, no standard procedure has been accepted. We understand that investigations on methodology are progressing both in the United States and in collaborative studies based in England as well. Our provisional table on carbohydrate fractions will help clarify the state of knowledge of carbohydrates and fiber, but definitive information will not result until there is an agreed-on-methodological approach to this problem.

Iodine is another nutrient of continuing interest, both from the standpoint of those who are concerned about meeting the requirement and those who are concerned that some people may be consuming an excess amount. Not only is the determination clouded by the lack of methods of proven reliability, but the distribution is affected by so many facets. The iodine content of plant foods is determined by the soil content, and this varies according to geographical region, which the iodine content of animal products is dependent upon the feeding practices employed. Milk is known to have a highly variable iodine content, not only because of variability in the feed but also as a result of the variable use of iodophor containing sanitizing agents at the farm and in the dairy. Bread is another product which may or may not contain relatively high amounts of iodine. This depends upon whether or not iodate is added in its preparation, either as a dough improver or as an ingredient of the yeast food. Thus, even if reliable methods were established, it would be difficult to compile meaningful, representative iodine values for foods consumed in the United States. As with selenium, iodine is a nutrient that will require special consideration in developing data on its distribution.

The problems remaining loom large and our rate of progress in defining the nutrient composition of foods still seems painfully slow. We remind ourselves that at the present time there is a greater amount of effort being applied to the nutrient analysis of foods than ever before in history. Although the goal may still be in the future, we can see measurable progress.

TABLE 1.--PUBLISHED SECTIONS OF AGRICULTURE HANDBOOK NO. 8

Number	Food group
8-1	Dairy and egg products
8-2	Spices and herbs
8-3	Baby foods
8-4	Fats and oils
8-5	Poultry products
8-6	Soups, sauces, and gravies
8-7	Sausages and luncheon meats
8-8	Breakfast cereals
8-9	Fruits and fruit juices

TABLE 2.--STATUS OF UNPUBLISHED SECTIONS OF AGRICULTURE NO. 8

Number	Food group	Preparation stage
8-10	Pork products	In press
8-11	Vegetables	In final review
	Nuts and seeds	DB III, exc. FA in DB I; AA in entry/correction
	Fish and shellfish	DB III, exc. AA in entry/correction
	Lamb, veal, game	Data entry/correction
	Beverages	DB II
	Legumes	DB II, exc. FA and AA in DB I (additional data expected)
	Cereal grains	Data collection and entry/correction, exc. FA in DB I
	Baked products	DB II, exc. FA in DB I; AA in entry/correction (additional data expected)
	Beef	Data entry/correction, exc. AA in DB I
	Sugars and sweets	Data collection and entry/ correction
	Mixed dishes	Data collection and entry/ correction
	Fast foods	DB II (additional data expected)
	Miscellaneous	Data collection

TABLE 3.--RECENT PROVISIONAL TABLES

Iron Content of Food (HERR 45)
Table of Amino Acids in Fruits and Vegetables
Provisional Table on Percent Retention of Nutrients in Food Preparation

Table 4

STATE OF KNOWLEDGE OF NUTRIENT COMPOSITION

	CARBOHYDRATES			LIPIDS					MINERALS										
	INDIVIDUAL SUGARS	STARCH	DIETARY FIBER	TOTAL FAT	FATTY ACIDS	CHOLES-TEROL	OTHER STEROLS	TRANS-FATTY ACIDS	CAL-CIUM	IRON	PHOS-PHORUS	SODIUM	MAGNES-IUM	POTASS-IUM	ZINC	COPPER	MANGAN-ESE	CHROMIUM IODINE SULFUR SELENIUM	COBALT, FLUORINE, NICKEL, SILICON, TIN, VANADIUM
BABY FOODS	--	--	I	S	I	--	--	--	S	S	S	S	I	S	I	I	--	--	--
BAKED FOODS, BREAD	--✓	I✓	I•✓	I•	I	--	--	I	S•	S•	I•	I•	I•	I•	I•	I•	--•	--	--
COOKIES, CRACKERS	--✓	I✓	I✓	I•	I	--	--	I	I•	I•	I•	I•	I•	I•	I•	I•	--•	--	--
SWEET GOODS	--✓	I✓	I✓	I•	I	--	--	I	I•	I•	I•	I•	I•	I•	I•	I•	--•	--	--
BEVERAGES	I•	NA	NA	--•	--	NA	NA	NA	I•	I•	I•	I•	I•	I•	I•	I•	--•	--	--
BREAKFAST CEREALS	S	I	S	S	I	--	--	NA	S	I	S	S	S	S	S	S	I	--	--
CANDIES	I•	--•	--•	I•	I	I	I	I	I•	I•	I•	I•	I•	I•	I•	I•	--•	--	--
CEREAL GRAINS, WHOLE	I•	I•	I•	I•	S	NA	--	NA	I•	I•	I•	I•	I•	I•	I•	I•	I•	I	I
FLOURS	I•	I•	I•	I•	S	NA	--	NA	I•	I•	I•	I•	I•	I•	I•	I•	--•	--	--
PASTA	I•	I•	I•	I•	I	NA•	--	NA	I•	I•	I•	I•	I•	I•	I•	I•	--•	--	--
DAIRY PRODUCTS	S	NA	NA	S	S	S	NA	NA	S	S	S	S	S	S	S	S	I	--	--
EGGS & EGG PRODUCTS	NA	NA	NA	S	S	S	NA	NA	S	S	S	S	I	S	S	I	I	--	--
FATS & OILS	NA	NA	NA	S	S	S	I	S	--	--	--	--	--	--	--	--	--	--	--
FISH & SHELLFISH, RAW	NA	NA	NA	S•	S•	S•	I•	--	S•	S•	S•	S•	S•	S•	S•	S	S	--	--
COOKED	NA	NA	NA	S	S	S	--	--	--	--	--	--	--	--	--	--	--	--	--
FRUITS, RAW	S✓	--✓	I•	S•	I	NA	I	NA	S•	S•	S•	S•	S•	S•	I•	I•	I•	--•	--•
COOKED, DRIED	I✓	--✓	I✓	--✓	--	NA	--	NA	--✓	--✓	--✓	--✓	--✓	--✓	--✓	--✓	--	--	--
FROZEN, CANNED	I✓	--✓	I✓	S✓	--	NA	--	NA	S✓	S✓	S✓	S✓	S✓	S✓	I✓	I✓	--✓	--	--
LEGUMES, RAW	I✓	I•	I•	S•	I	NA	I	NA	S•	S•	S•	S•	S•	S•	S•	S•	S•	--	--
COOKED	I✓	I•	I•	S•	I	NA	--	NA	S•	S•	S•	S•	S•	S•	S•	S•	S•	--	--
PROCESSED	I✓	I✓	I✓	I✓	I	NA	--	NA	I✓	I✓	I✓	I✓	I	I	--	--	--	--	--
MEAT, RAW & COOKED, BEEF	NA	NA	NA	S	S	S	NA	NA	S	S	S	S	S	S	S	S	S	--	--
LAMB	NA	NA	NA	S	S	S	NA	NA	S	S	S	S	S	S	S	S	S	--	--
PORK, FRESH & CURED	NA	NA	NA	S	S	S	NA	NA	S	S	I	S	S	S	S	S	S	--	--
VEAL	NA	NA	NA	S	S	S	NA	NA	S	S	S	S	S	S	S	S	S	--	--
SAUSAGE & LUNCH MEAT	NA	NA	NA	S	S	S	NA	NA	S	S	S	S	S	S	S	S	S	--	--
NUTS & SEEDS	I•	I•	I•	S•	S	NA	I	NA	I•	I•	I•	I•	I•	I•	I•	I•	I•	I	I
POULTRY, RAW & COOKED	NA	NA	NA	S	S✓	S	NA	NA	S✓	S✓	S✓	S✓	S✓	S✓	S✓	S✓	S✓	--✓	--✓
SNACK FOODS	--	--	--	--	I	--	--	I	--	--	--	--	--	--	--	--	--	--	--
SOUPS	--	--	--	S	I	I	I	--	S	S	S	S	S	S	S	S	--	--	--
VEGETABLES, RAW	I•	S•	I•	I•	I	NA	I	NA	I•	I•	I•	I•	I•	I•	I•	I•	--•	--•	--•
COOKED	I•	I•	I•	--•	--	NA	--	NA	--•	--•	--•	--•	--•	--•	--•	--•	--	--	--
FROZEN	I	I	I	S	--	NA	--	NA	S	S	S	S	S	S	S	S	--	--	--
CANNED	I✓	I✓	I✓	S✓	--	NA	--	NA	S✓	S✓	S✓	S✓	S✓	S✓	S✓	S✓	--✓	--	--
FAST FOODS	--•	--•	I•	S•	I•	I•	I	I	S•	S•	S•	S•	S•	S•	S•	S•	--•	--	--
FROZEN DINNERS	--	--	I	I	I	--	--	--	I	I	I	I	I	I	--	--	--	--	--
INSTITUTIONAL FOOD	I	I	I	I	I	I	--	--	--	--	--	--	--	--	--	--	--	--	--
MIXED DISHES, COMMERCIAL	--	--	--•	I•	I	I•	I	--	S•	S•	S•	S•	S•	S•	I•	I•	--•	--	--
HOME PREPARED	--	--	--•	I•	I	I•	--	--	I•	I•	I•	I•	I•	I•	--•	--•	--•	--	--
RESTAURANT FOOD	I	I	I	--•	--	--	--	--	--•	--•	--•	--•	--•	--•	--•	--•	--	--	--

-- LITTLE OR NO DATA
S SUBSTANTIAL DATA
I INADEQUATE DATA
NA NOT APPLICABLE
• USDA/CND CONTRACTS IN PROGRESS
✓ USDA/CND CONTRACTS PROPOSED FOR FISCAL 1983

NUTRIENT DATA RESEARCH BRANCH
CONSUMER NUTRITION DIVISION
HUMAN NUTRITION INFORMATION SERVICE
U.S. DEPARTMENT OF AGRICULTURE
HYATTSVILLE, MARYLAND 20782
JANUARY 1983

Table 4 (continued)

STATE OF KNOWLEDGE OF NUTRIENT COMPOSITION

PROTEIN AND AMINO ACIDS [1]

	TOTAL PROTEIN	CYSTINE METHIONINE	TRYPTOPHAN	ISOLEUCINE LEUCINE VALINE	SERINE THREONINE TYROSINE	ALANINE, ARGININE, ASPARTIC ACID, GLUTAMIC ACID, GLYCINE, HISTIDINE, LYSINE, PHENYLALANINE, PROLINE
BABY FOODS	S	I	I	I	I	I
BAKED FOODS, BREAD	S•	I✓	I✓	S✓	S✓	S✓
COOKIES, CRACKERS	S•	I✓	I✓	I✓	I✓	I✓
SWEET GOODS	S•	I✓	I✓	I✓	I✓	I✓
BEVERAGES	I•	NA	NA	NA	NA	NA
BREAKFAST CEREALS	S	I	I	I	I	I
CANDIES	I•	--	--	--	--	--
CEREAL GRAINS, WHOLE	S•	I	S	S	S	S
FLOURS	S•	I	S	S	S	S
PASTA	S•	I✓	S✓	S✓	S✓	S✓
DAIRY PRODUCTS	S	S	S	S	S	S
EGGS & EGG PRODUCTS	S	S	S	S	S	S
FATS & OILS	NA	NA	NA	NA	NA	NA
FISH & SHELLFISH, RAW	S	S✓	S✓	S✓	S✓	S✓
COOKED	I	I	I	I	I	I
FRUITS, RAW	S•	I✓	I✓	I✓	I✓	I✓
COOKED, DRIED	I	I	I	I	I	I
FROZEN, CANNED	S	I	I	I	I	I
LEGUMES, RAW	S•	S	S	S	S	S
COOKED	S•	I✓	I✓	I✓	I✓	I✓
PROCESSED	S	I	I	I	I	I
MEAT, RAW & COOKED, BEEF	S	S	S	S	S	S
LAMB	S	S	S	S	S	S
PORK, FRESH & CURED	S	S	S	S	S	S
VEAL	S	S	S	S	S	S
SAUSAGE & LUNCH MEAT	S	S	S	S	S	S
NUTS & SEEDS	S•	I✓	I✓	I✓	I✓	I✓
POULTRY, RAW & COOKED	S✓	S	S	S	S	S✓
SNACK FOODS	--	--	--	--	--	--
SOUPS	S	I	I	I	I	I
VEGETABLES, RAW	S•	I	I	S	S	S
COOKED	S•	I	I	I	I	I
FROZEN	S	I	I	I	I	I
CANNED	S	I	I	I	I	I
FAST FOODS	S	I•✓	I•✓	I•✓	I•✓	I•✓
FROZEN DINNERS	I	--✓	--✓	--✓	--✓	--✓
INSTITUTIONAL FOOD	--	--•✓	--•✓	--•✓	--•✓	--•✓
MIXED DISHES, COMMERCIAL	S	--✓	--✓	--✓	--✓	--✓
HOME PREPARED	I•	--✓	--✓	--✓	--✓	--✓
RESTAURANT FOOD	--•	--	--	--	--	--

VITAMINS

	VITAMIN A	THIAMIN	RIBOFLAVIN	VITAMIN B_6	VITAMIN B_{12}	ASCORBIC ACID	NIACIN	FOLACIN	VITAMIN D	VITAMIN E	BIOTIN	CHOLINE	PANTOTHENIC ACID
BABY FOODS	S	S	S	I	I	S	S	I	NA	I	--	--	--
BAKED FOODS, BREAD	NA	S•	S•	I•	NA	NA	S•	I•	NA	I	--	--	I•
COOKIES, CRACKERS	--	I•	I•	I•	--	NA	I•	I•	NA	I	--	--	--•
SWEET GOODS	--	I•	I•	I•	--	NA	I•	I•	NA	I	--	--	--•
BEVERAGES	I•	I•	I•	I•	NA	I•	I•	I•	NA	NA	NA	NA	I•
BREAKFAST CEREALS	I•	I•	I•	I•	I	I•	I•	I•	--	I	--	--	I•
CANDIES	NA	I•	I•	I•	NA	I•	I•	I•	NA	--	--	--	I•
CEREAL GRAINS, WHOLE	I	I•	I•	I•	NA	I	I•	I•	NA	I	--	S	I•
FLOURS	--•	I•	I•	I•	NA	NA	I•	I•	NA	I	--	I	I•
PASTA	--•	I•	I•	I•	NA	NA	I•	I•	NA	--	--	--	--•
DAIRY PRODUCTS	S✓	S	S	S	S	S	S	I	S	I	--	I	S
EGGS & EGG PRODUCTS	I✓	S	S	S	S	NA	S	I	I	I	--	S	S
FATS & OILS	--	--	--	--	--	--	--	--	--	--	--	I	--
FISH & SHELLFISH, RAW	I•	I•	I•	I•	I•	NA	I•	--•	I	I	--	I	I•
COOKED	--	--	--	--	--	--	--	--	I	I	--	--	--
FRUITS, RAW	S•	S•	S•	S•	NA	S•	S•	I•	NA	--	--	I	I•
COOKED, DRIED	I	I	I	I	NA	I	I	I	NA	--	--	--	I
FROZEN, CANNED	S	S	S	I	NA	S	S	I	NA	--	--	--	I
LEGUMES, RAW	--	S•	S•	I•	NA	NA	S•	I•	NA	S	--	I	I•
COOKED	--	S•	S•	I•	NA	NA	S•	I•	NA	I	--	--	I•
PROCESSED	I	I✓	I✓	--✓	NA	I✓	I✓	--✓	NA	I	--	--	--
MEAT, RAW & COOKED, BEEF	S	S	S	S	S	NA	S	S	I	I	--	I	I
LAMB	I	S	S	S	S	NA	S	S	I	I	--	--	I
PORK, FRESH & CURED	I	S	S	S	S	NA	S	I	I	I	--	I	I
VEAL	I	S	S	S	S	NA	S	S	I	I	--	--	I
SAUSAGE & LUNCH MEAT	I	S	S	S	S	S	S	I	I	--	--	--	I
NUTS & SEEDS	I	I•	I•	I•	NA	I	I•	I•	NA	I	I	I	I•
POULTRY, RAW & COOKED	I✓	S✓	S✓	I✓	I✓	NA	S✓	I✓	I	I	--	I	I✓
SNACK FOODS	--	--	--	--	--	--	--	--	--	--	--	--	--
SOUPS	S	S	S	I	I	S	S	--	NA	--	--	--	--
VEGETABLES, RAW	S•	S•	S•	I•	NA	S•	S•	I•	NA	I	--	I	--•
COOKED	I•	I•	I•	I•	NA	I•	I•	I•	NA	I	--	--	--•
FROZEN	S	S	S	--	NA	S	S	I	NA	--	--	--	I
CANNED	S✓	S✓	S✓	--✓	NA	S✓	S✓	I✓	NA	--	--	--	I✓
FAST FOODS	S•	S•	S•	I•	I•	I•	S•	--•	--	I	--	--	--
FROZEN DINNERS	I	S	S	--	--	I	S	--	--	--	--	--	--
INSTITUTIONAL FOOD	--	--	--	--	--	--	--	--	--	--	--	--	--
MIXED DISHES, COMMERCIAL	I	S	S	I	I	I	S	I	--	I	--	--	--•
HOME PREPARED	I•	I•	I•	--•	--•	--•	I•	--•	--	--	--	--	--•
RESTAURANT FOOD	--•	--•	--•	--•	--	--•	--•	--•	--	--	--	--	--•

-- LITTLE OR NO DATA
S SUBSTANTIAL DATA
I INADEQUATE DATA
NA NOT APPLICABLE
• USDA/CND CONTRACTS IN PROGRESS
✓ USDA/CND CONTRACTS PROPOSED FOR FISCAL 1983

NUTRIENT DATA RESEARCH BRANCH
CONSUMER NUTRITION DIVISION
HUMAN NUTRITION INFORMATION SERVICE
U.S. DEPARTMENT OF AGRICULTURE
HYATTSVILLE, MARYLAND 20782
JANUARY 1983

Table 5

STATE OF DEVELOPMENT OF METHODS FOR NUTRIENTS IN FOODS

Nutrient Composition Laboratory
ARS, USDA
Beltsville, MD 20705
July 1983

	State of Methodology[a/]			
Nutrient category	Adequate	Substantial	Conflicting	Lacking
Carbohydrates, fiber and sugars		Individual sugars	Fiber Starch	
Energy			Food energy	
Lipids		Cholesterol Fat (total) Fatty acids (common)	Sterols Trans-fatty acids	
Minerals/Inorganic nutrients	Calcium Copper Magnesium Phosphorus Potassium Sodium Zinc	Iron (total) Selenium	Arsenic Chromium Fluorine Iodine Manganese	Cobalt Heme-iron Molybdenum Nonheme-iron Silicon Tin Vanadium
Proteins and amino acids	Nitrogen (total)	Amino acids (most)	Amino acids (some) Protein (total)	
Vitamins		Niacin Riboflavin Thiamin Vitamin B-6	Vitamin A Carotenes Vitamin B-12 Vitamin C Vitamin D Vitamin E Folacin Pantothenic acid	Biotin Choline Vitamin K

a/ Description of methodology states

Factors	Adequate	Substantial	Conflicting	Lacking
Accuracy	Excellent	Good	Fair	Poor
Speed of analysis	Fast	Moderate	Slow	Slow
Cost per analysis	Modest (<$100)	Modest to high	High	?
Development needs	—	Method modif. Extraction proc. Applications	Method develop. modif. Extraction proc. Applications	Method develop. Extraction proc. Applications

From: Beecher, G.R. and Vanderslice, J.T. Determination of Nutrients in Foods. In: Proceedings of Symposium on Modern Methods of Food Analysis. Editors, K.K. Stewart and J.R. Whitaker. Westport: AVI, 1984 (1984).

NUTRIENT DATA BASE FOR A SURVEY OF INDIVIDUAL INTAKES: NATIONWIDE FOOD CONSUMPTION SURVEY 1977-78

Eleanor M. Pao
Human Nutrition Information Service
U.S. Department of Agriculture

Large-scale surveys such as the Nationwide Food Consumption Survey (NFCS) conducted by the U.S. Department of Agriculture (USDA) from April 1977 through March 1978 require a support system that includes several components. They are:

- names and descriptions of foods,
- a system for coding food items,
- conversions for common measures to gram weights (of quantities reported),
- rules for handling inadequately reported information, and
- food composition values for energy and nutrients.

These components are essential for processing the responses obtained in the survey in order to meet survey objectives. They made up the structure of the food/nutrient data base for the survey of individual intakes as used in the NFCS 1977-78. In the NFCS two types of food information were collected--food as used by the household unit and food as ingested by individual members. This paper deals only with the survey of food intakes by individuals. All parts of the food/nutrient data system will be discussed.

Individuals generally reported food items on the questionnaires (a 1-day recall and a 2-day record) in the order in which they were eaten, starting with the first meal or snack of the day and ending with the last. Respondents were not asked to categorize foods in any way other than by time of day each eating occasion began. This was regarded as the best way to get nonconditioned response. Thus, although data were collected in such a way that food items could be identified and classified under specific codes, the data were essentially unstructured. The coding system served as a tool to organize the individual responses for the initial processing steps such as application of the nutrient composition file and later in the execution of tabulations and analyses. Hence, the coding system gives structure which is lacking in the individual intake questionnaire. (This is in contrast to the household food use questionnaire in which foods are organized into categories and precoded.)

The content and organization of a data base are governed by objectives of the survey, characteristics of the survey responses, and plans for use of the data base. Factors considered in organizing the NFCS 1977-78 data base included:

- nutrients of interest for which there were suitable food composition data;
- nonnutritive substances of current or likely future concern (such as caffeine);
- quality and detail of collected food intake information;
- types of tabulations and analyses to be given priority;
- maintenance of linkage to data from the 1965 survey so as to study changes or trends;

- nutrient information to be retained in the basic code and nonnutrient information--such as packaging--that could be retained as a separate variable;
- minimum length of code to store information on factors affecting nutrients yet retaining flexibility for insertion of new codes;
- use of recipes or complete mixtures versus total or partial separation into ingredients or component parts;
- use of brand names versus generic identification only, as for fast foods, TV dinners, ready-to-eat cereals, and candies;
- availability of information on enriched and fortified products versus nonenriched and nonfortified products.

FOOD INTAKES--DATA COLLECTION AND DATA REDUCTION

The quality of responses in a survey materially affects how well the data base or support system will perform, and the quality of responses obtained in surveys is affected by data collection procedures. Respondents must be able and willing to give the information requested; otherwise, survey results will be invalid. The burden on the respondents must be within reasonable limits to prevent loss of interest and fatigue which can lead to poor recall and careless reporting. After data collection, appropriate treatment of the responses by reviewers/editors and coders is vital to the data reduction process in order to retain the integrity of the information about respondents' intakes.

NFCS data were collected by previously experienced interviewers who received 1 week of training in the NFCS methods. After administration of the household food use phase of the questionnaire, the interviewer administered the 1-day food intake recall questionnaire to each eligible household member present. (The day began at midnight, 12:00 a.m., and ended at 11:59 p.m.) Then individuals received instructions for keeping a record of foods and beverages consumed during the day of and the day following the interview. A set of stainless steel measuring cups and spoons and a plastic ruler was given to each household to help estimate portion sizes. A leaflet on how to describe foods and amounts eaten was left with each respondent. The interviewer coached individuals as they recorded the foods and beverages eaten earlier on the interview day. The household respondent usually reported for children under 12 and others unable to answer for themselves. Forms were left for absent members to fill out. Interviewers arranged for a return visit to review and pick up the 2-day records. Households were paid $1 for each completed 2-day record up to a limit of $10 per household.

NFCS respondents were instructed to describe each food or beverage item in detail--kind, form, preparation, brand name if well known, and any other relevant information. Each food item was recorded on a separate line along with the description and quantity eaten. The description was supposed to indicate details regarding presence of refuse in the quantity reported. Quantities were reported in measures respondents found most convenient. However, respondents were encouraged to use the measuring cups and spoons in measuring fluids and foods commonly measured by cups and spoons. Foods served by the piece were often reported in dimensions or as number of pieces. Package weight was often used for foods that were purchased in packages with weights printed on labels. If the food was a steak, for example, the weight might be for raw meat and it could include bone and uneaten fat.

The description was expected to indicate the nature of the amount reported so that the yield as ingested (cooked, without bone and uneaten fat) could be computed. Interviewers questioned respondents about entries for items suspected to include refuse along with the amounts ingested. Such probing by interviewers enhances the quality of dietary survey data.

Questionnaires were reviewed in the central office. Reviewers/editors checked for clarity, consistency, completeness, and reasonableness and decided if callbacks to the field were necessary. If not, the questionnaires were forwarded to the coders. Coding of food intake information was entered directly onto the questionnaires by trained coders following the coding manual. All food-item codes and amounts were then checked by supervisors. Guidelines were provided to achieve uniformity in the coding operation. The coded food-intake data were keyed and transferred to computer tape and thus became input into the food intake data base for subsequent processing.

The coding manual contained the several types of information which coders needed to code food intakes. These were (1) the name and description of each food item, (2) the code for each food item, (3) common measures of portions as reported, (4) weights in grams of common measures as ingested, and (5) default values to use for imprecisely reported portion sizes.

CONCEPTUALIZATION OF THE CODING SYSTEM

The NFCS coding system was planned to organize foods reported in the survey into a data base that would meet needs of prospective users of the survey results. A major principle guiding formation of the coding system was establishment of several levels of subcategories within major food groups to form hierarchies and permit regrouping on different dimensions to meet special needs. The subgroups could also be collapsed to yield more highly aggregated data. Another principle was that the coding system be flexible so that new codes could be added without disrupting existing food groupings.

A third principle was that the codes for food items should permit extraction of information of nutritional importance. A separate variable would be created to provide ancillary information such as whether the food was a single or multiple component item, and, if multiple, what it was, e.g., a sandwich made up of bread and filling or a salad with ingredients listed separately. Two additional principles:were related to this third principle. The system required that each food item be uniquely identified by its code and be described in sufficient detail that appropriate energy and nutrient values could be assigned. Also, the length of the code should be kept short so as to minimize error. A food code with seven digits was selected. Because people are accustomed to using seven-digit telephone numbers, this length of a code was viewed as practical. It was also decided, after NFCS was underway, that new foods would be added to the data base only if they were reported by respondents.

After identification of principles to guide the organization of the coding system, criteria for classifying foods into subgroups were specified. As the level of subgrouping becomes more detailed, the foods in each subgroup become more homogeneous. Among the criteria underlying formation of food groupings were the following:

- abundant source of a particular nutrient such as calcium in milk and vitamin A in deep yellow vegetables;
- varieties of a food such as varieties of cheese;
- common stage of preparation such as raw vegetables;
- common method of processing such as dried fruits;
- common method of cooking such as baked white potatoes;
- a common form of the food such as juice of citrus fruits;
- preparation or use largely for a particular age group such as commercial baby foods and baby formulas;
- accessory role of items such as pickles, relishes, olives, and table fats;
- traditional role in meals such as desserts (milk desserts);
- similarity in supplying mainly energy in the diet such as fats and oils;
- common usage of an item such as salty snacks;
- imitation, substitute, or formulated foods intended to replace a natural food such as meatless meats or soy-based milk;
- addition of a particular nutrient such as vitamin C to fruit drinks;
- modification of energy content of a type of food such as low-calorie salad dressings and diet beverages;
- common type of mixture such as soups;
- mixtures comprised of several common components such as beef with starch and vegetables;
- presence of a major common component such as alcohol in alcoholic beverages;
- a common cut of meat or poultry such as steaks or chicken breast; and
- presence of inedible parts in portions served that require computation of yield.

OPERATIONALIZING THE CODING SYSTEM

Several of the above criteria were applied in selecting the key categories of the coding system. Nine basic food commodity groups were identified by the first digit in the seven-digit food code. They were:

(1) milk and milk products;
(2) meat, poultry, and fish;
(3) eggs;
(4) legumes, nuts, and seeds;
(5) grain products;
(6) fruits;
(7) vegetables;
(8) fats and oils; and
(9) sweets, sugars, and beverages.

A tenth category, identified by 0, was created to include substances for which no nutritive values were prided in the nutrient data base. Substances such as flavorings, seasonings, and nonfood supplements were placed in this category.

The next two digits in the food code organized foods into subgroups formulated in accordance with the criteria just noted. The second digit identified major subgroupings within commodity or major food groups, and the third digit sorted foods in the major subgroupings into minor subgroups.

For example, the second digit organized fruits into citrus fruits and juices (1), dried fruits (2), fruits and berries excluding citrus and dried (3), noncitrus fruit juices and nectars (4), and baby-food fruits and juices (7); the third digit organized citrus fruits and juices into citrus fruits (1) and citrus fruit juices (2). The fourth and fifth digits were also used for this purpose in large food groups as, for example, in the fruits group. The sixth and seventh digits were reserved to carry information unique to a food item; for example, for peaches (coded 631-3562) the sixth digit indicated "frozen" and the seventh digit indicated "unsweetened."

The types of information coded in the fourth through seventh positions varied from one food subgroup to another in order to reflect the characteristics peculiar to each subgroup and depending on the size of the food group. The use of the digits in the food code to subgroup and to provide unique identification is illustrated by the fruit group's codes in Table 1.

Generally, the more nonzero digits in the fourth through seventh positions, the more detailed the description is. The more zeros in these positions, the more general is the description provided by the code. To denote the absence of detailed information, such a food within a food group was assigned a code defined as NFS for "Not Further Specified." Each NFS food item was usually assigned an identity based on the form of the food that was most frequently consumed, had the largest market share, or was representative of several such items. Respondents might not be able to describe adequately foods eaten away from home, but they could usually provide enough information to code at least the first two digits.

Information on new or unusual foods was obtained by visiting or calling stores and restaurants; from food industry fact sheets, newspapers, magazines, and cookbooks; by phone calls to foreign embassies, Government agencies such as the National Marine Fisheries Administration for fish and the Fish and Wildlife Service for game. Information on school lunch foods was found in the list of menu items supplied by the Food and Nutrition Service. Cooperative Extension Service agents provided information about unusual local foods. Nutritionists of particular ethnic backgrounds, for example, Puerto Rican, were also valuable resource people.

Some foods could be organized into more than one subgroup, depending upon the common element selected. Because only one element could be used, some arbitrary decisions about grouping foods had to be made. Mixtures were assigned to the food groups of the primary component or ingredient. Meat was usually considered to be a primary component, but sometimes it was not the largest ingredient in quantitative terms. If only small amounts of meat were included, the mixture was assigned to the food group of the major ingredient. For example, pizza was put in the grain products group. Insofar as possible, recipes were used for mixtures. But there were also instances where the components of a mixture were coded individually, especially if reported separately. If this was done for a sandwich, frozen meal, or salad, the action was documented in a separate variable (or special code) that specified the kind of mixture. Information on salads put together at salad bars could often be more accurately represented by coding each item separately. Sandwiches purchased as a unit were usually reported and coded that way but if the parts were reported separately, they were coded separately.

There are advantages to having basic food items and their "add-ons" coded individually because then they can be assigned to the most appropriate food groups and in the amounts actually eaten.

Codes for individual brand name products were used for a few products, but most food codes represented a more broadly based item, e.g., several very similar brands or generic foods. Ready-to-eat cereals were an exception, however, because formulations were often unique for each cereal, and many were highly fortified. In such cases, brand name products were identified with unique codes, although the brand names do not appear on the data tapes. In other cases, brand name products were used as the sources of information for food items because they were representative of similar foods or because they were the best sources of information for particular food items.

Some decisions about whether a food item should be assigned a new food code or be included in a group of similar foods relied on professional judgment. In some cases, the decision was based on the possible benefit from retaining identity of the food item for later reference. If the food item was assigned the food code of an existing item, its identity as a separate item was lost. Liquid protein diet is a good example. There was great concern about use of this product. Unless it were given a unique code, it would be difficult to retrieve information on usage without going back to the original records.

Application of the food coding system was not free of problems. Descriptions of food items were sometimes more detailed than coders wished, especially when respondents supplied "too much" or irrelevant information. However, respondents were more likely to report too little descriptive information than too much. For this reason, NFS food item codes were used. Frequency of use of some well-defined foods, e.g., milk with 2 percent fat, probably understates actual usage because some respondents using the product probably reported only "milk." If so, the food item was assigned an NFS code.

More detailed information can be retained in a data base if the number of food codes is not limited in any way. But with a large number of food codes, coding is more difficult and tedious because the coder must consider more food codes before selecting the code that best matches the respondent's description of the food item.

FREQUENCIES OF CODES USED IN THE 1977-78 CODING SYSTEM

How well did the coding system work for NFCS 1977-78? The total number of food codes in the data base reached by the end of the processing of the basic and supplemental surveys was 4,546 and 25 codes which were discontinued during the course of the survey (total = 4,571). (Supplemental surveys included those in Hawaii, Alaska, Puerto Rico, and special surveys in the 48 states among low-income households and among elderly households.) The number of codes used in the basic survey in the 48 states during the year was 3,702. Thus, 844 of the total codes in the data base were not used in the basic survey. Some of those unused codes were for food items retained from the 1965 survey which were not reported in NFCS 1977-78 and others were used only in the supplemental surveys.

Of the 3,702 food codes used, only 208 (5.6 percent) were assigned 1,000 or more times and only 1,005 food codes were assigned 100 or more times (27.1 percent). (See Table 2.) Of the 3,702 codes, 1,249 (33.7 percent) were used less than 10 times and 281 (7.6 percent) were used but once. The grain products group had the most food codes that were used 1,000 times or more, for 53 items (6.3 percent of the codes used in the grain products group). Of all food groups in the data base, the meat, poultry, and fish group had the largest number of used food codes, 1,069 items, of which 217 (20.3 percent) were used 100 or more times. (When uses were weighted to maintain the representativeness of the survey sample, frequencies were slightly different.)

Not Further Specified (NFS) food codes accounted for 26.2 percent of all foods and beverages reported on questionnaires. There were 572 NFS codes in the data base (15.5 percent of total food codes). The number of NFS food item codes among the 208 food codes which were assigned 1,000 times or more was 62; this was 29.8 percent of those codes used 1,000 or more times. Among those food codes used fewer than 5 times, 61 were NFS food codes (7.8 percent of items used 1 to 4 times). A large number of the seldom-used codes were in the meat group, especially for chicken and fish because many food codes in those subgroups had detailed descriptions. (See Table 3.)

Of the total food items reported in the survey from all food groups, the most belonged to the grain products group, 21.3 percent of all items reported, and the sugar, sweets, and beverage group was next with 19.4 percent. (See Table 4.) Other food groups accounted for fewer items reported. In descending order they were: milk and milk products (15.3 percent), vegetables 15.1 percent), meat group (12.6 percent), fruits (6.5 percent), fats and oils (6.1 percent), eggs (2.2 percent), and legumes and nuts (1.7 percent). Of total reported food items, 73.9 percent were accounted for by the 208 food codes used 1,000 or more times. The meat and legumes groups had the smallest proportion of their total reported items in the list of food codes used 1,000 or more times--49.9 percent and 58.6 percent, respectively.

Figures like those just mentioned provide one basis for determining how well the data base performed and how it might be revised for the next survey. Food items that were seldom reported need to be examined to determine whether they should be retained or are sufficiently similar to other food items to be combined with them. However, some food codes will need to be retained despite infrequent usage so that future usage by the population can be measured, e.g., carob.

CONVERSIONS OF MEASURES TO GRAM WEIGHTS

Quantities of food and beverages were reported as ingested in whatever way was most convenient for the respondent, but all were finally converted to the common unit, grams, for analysis. Measures commonly used in reporting each food item were listed in the coding manual along with the weight of each measure in grams. The set of standard measuring cups and spoons and the plastic ruler were expected to aid respondents in estimating amounts ingested.

"Edible portions" and portions "as ingested" can differ in a number of foods, such as meats. An edible portion of meat can be raw; whereas a portion "as ingested" is usually cooked. Also, a portion "as served" may include refuse and differ from the portion "as ingested." If amounts reported included refuse, such as rind or bones, or the amounts were for a food before cooking, e.g., raw steak or a frozen package of spinach, the yield for the amount ingested was provided. The source of information for calculating yields was Agriculture Handbook No. 102. The most frequent source of weights for common measures was Agriculture Handbook No. 456. For new foods, weights of common measures often had to be obtained by store checks, telephone calls to fast food chains or other business firms, food product fact sheets, or by purchasing and actually measuring and weighing the product.

The use of dimensions to describe regularly shaped pieces or portions was usually satisfactory if measured carefully. In such cases, the number of cubic inches in a portion was computed and then converted to grams by means of weight per cubic inch. Such factors are still tentative for a number of food items in the data base. Whenever imprecise terms such as "small," "medium," and "large" were reported, specifications of measurements regarded as small, medium, or large were necessary.

For respondents' dietary records to be acceptable, the guidelines stipulated that at least 9 out of 10 foods had to have amounts reported in measures readily converted to grams. In case portions were inadequately reported (such as "one serving") or not reported at all, a default value called "NS" (Not Specified) was provided. Every use of a default value was documented by use of a special code.

The common measures and the equivalent weight in grams for each food item coded in the data base made up a "gram conversion file" on computer tape. This file was used to compute the gram weights of measures not already in grams. Subsequently, the grams ingested of a food item were used to calculate the food's energy and nutrient content.

ENERGY AND NUTRIENT VALUES IN THE DATA BASE

The nutrient data base for NFCS 1977-78 included food composition values for energy and 14 nutrients per 100 grams for every food code in the data base. The 14 nutrients (and their units of measure) were protein (g), fat (g), carbohydrate (g), calcium (mg), iron (mg), magnesium (mg), phosphorus (mg), vitamin A value (IU), thiamin (mg), riboflavin (mg), preformed niacin (mg), vitamin B6 (mg), vitamin B12 (mcg), and vitamin C (mg). Those energy and nutrient values were supplied by the USDA Nutrient Data Research Branch (NDRB) of the Consumer Nutrition Division. Agriculture Handbook No. 8 and three of its revisions (Nos. 1, 3, and 5) were the basic source of values in the NFCS 1977-78 nutrient data base. Data supplied by food industries were also used. The source of vitamin B6 and vitamin B12 values was Home Economics Research Report No. 36.

Since the data base for individual intakes used in the Household Food Consumption Survey conducted in 1965 was the core for the NFCS 1977-78 data base, all energy and nutrient values were reviewed and most were updated to reflect recent data. The enrichment standards enacted in 1975 for bread and flour were reflected in the updated values for thiamin, riboflavin, and niacin. As new foods were reported in the survey, they were added to the nutrient data base with appropriate values assigned.

As indicated in discussing foods and their descriptions, inadequately described foods (Not Further Specified) were assigned values based on one or more of the most commonly used items in the food subgroup. Energy and nutrient values for recipes were usually based on recipes from popular cookbooks.

For the limited number of brand name products that appear in the data base, the energy and nutrient values are unique to the individual food item. In most cases, the energy and nutrient values for food items were broader and encompassed consideration of a variety of brands or food from a variety of geographical areas, seasons, stages of maturity, and other factors that affect food composition values.

A number of different food items in the data base have the same energy and nutrient values assigned to them. This happened if some attribute other than nutrient value was being tracked--such as species of fish for toxic contamination--or in the case of coffee, whether it was decaffinated or not. The forms in which foods were marketed--such as cuts of meat (ground versus solid cut)--might have the same nutrient values but their usage was important to know.

A calcium equivalent factor was also included in the nutrient data base for foods in the milk and milk products group. It was an expression (in grams) of the amount of fluid whole cow's milk that has the same quantity of calcium as the reported food. This factor provided a basis for aggregating different forms of milk and milk products into one amount.

UPDATING THE SUPPORT SYSTEM FOR THE NEXT SURVEY

To meet the needs of users of data from the next nationwide survey of individuals' diets, the several types of data in the support system will have to be updated. Some types of desirable updating are already evident from the experiences in analyzing the 1977-78 data and in responding to the large volume of requests from private industry, nutrition and health professionals, researchers, private and government agencies, the media, and many others.

There have been repeated requests for information on nutrients not covered in 1977-78--zinc, sodium, potassium, folacin, and others--as well as on substances such as sugars and fiber. The feasibility of adding nutrients is under consideration. The limitations of available analytical data must not be overlooked. Obviously new foods will be added to the data base--such as low-sodium and diet foods--as their prominence in the marketplace increases.

Users have also expressed interest in having more nonnutrient information about food items such as type of packaging, kind and amount of home preparation, how foods are used in a meal, and brand names. Such information could be retained in separate variables similar to the variable on whether the food is a single or multiple unit.

There is a need for careful review and revision of the data base with the objective of reducing or combining a number of similar food items in several food groups. Seldom-used food codes will be reviewed to determine whether they should be dropped or handled in another manner. However, this has to be balanced against the need for detailed information. Default values (in grams) and guidelines for handling inadequate information also need review.

Some of the criteria to be used in reviewing the current data base of food items include the following:

- Is the food item still used?
- Is the form still sufficiently important to be coded separately?
- Does the identity of the food item need to be retained to show extent of use by the population?
- Are separations of food items grouped under one code necessary if additional nutrients are included in future surveys?
- Can separation of a component (such as "fat added" for vegetables) simplify and reduce necessary codes without losing information?
- Can refuse such as bones and shells be handled as a yield measure and therefore codes "with bone," "with shell," etc., be deleted?
- Are "homemade" items sufficiently used and nutrient values sufficiently different to justify separate codes?
- Should nonnutrient information be carried as a separate variable for food items?
- Should some food items be moved to another food group such as soy sauce from nonnutrient group to soy products group?
- Should brand name foods such as candies be added to data base?

Needless to say, careful planning and preparation of the data base or support system for a survey is of great importance. A well-organized data base contributes to timely release of results and simplifies meeting requests for special information.

Table 1. Use of digits in the food code to subgroup and to uniquely identify items in the fruit group--food intakes of individuals, NFCS 1977-78

1 (Main category)	2 (Major subgroup)	3 (Minor subgroup)	4-5 (Further subgrouping)	6-7 Unique identity of item)
6 Fruits	61 Citrus fruits, juice	611 Citrus fruits		
		612 Citrus juices		
	62 Dried fruits	621 Dried fruits		
	63 Fruits and berries, exclude citrus and dried	631 Fruits, exclude berries	-01 Apples	
			-03 Apricots	
			--	
			-07 Bananas	
			--	
			-23 Grapes	
			--	
			-35 Peaches	3500 NFS
			--	3501 Raw, NFS
				3512 Cooked, canned, unsweetened
				3513 Cooked, canned, heavy sirup
				3514 Cooked, canned, light sirup
				3514 Cooked, canned, drained solids
				3516 Canned, juice pack
				3561 Frozen
				3562 Frozen, unsweetened
				3563 Frozen, sweetened
				3565 Pickled
			-37 Pears	
			--	
		632 Berries		
		--		
		634 Mixtures of fruits		
	64 Fruit juices, nectars	641 Juices		
		642 Nectars		
	67 Baby and junior fruits and juices	671 Fruits and fruit mixtures		
		--		
		674 Fruit desserts and puddings		

Dashes (--) indicate codes and food items in series are omitted because of space limitation.

Table 2. Number of food codes used designated number of times in Basic NFCS 1977-78, 3-day reports only, 48 States

Food group	Total codes used	Number of times food code was used: 1,000 and over	500-999	100-499	50-99	10-49	5-9	3-4	2	1	Under 100
		Number of food codes									
Milk and milk products..	280										
Unweighted...........		21	18	59	30	77	29	12	14	20	182
Weighted.............		31	17	62	38	75	20	18	11	8	
Meat group.............	1,069										
Unweighted...........		32	28	157	104	324	154	89	84	97	852
Weighted.............		43	44	182	112	341	138	93	70	46	
Eggs..................	45										
Unweighted...........		5	1	5	4	15	8	3	0	4	34
Weighted.............		5	2	5	6	18	5	0	1	3	
Legumes, nuts, seeds....	123										
Unweighted...........		4	3	23	10	36	20	8	11	8	93
Weighted.............		5	5	23	14	39	20	6	6	5	
Grain products..........	844										
Unweighted...........		53	47	181	97	228	85	70	35	48	563
Weighted.............		77	52	202	95	231	90	48	25	24	
Fruits..................	377										
Unweighted...........		14	13	49	35	113	56	37	21	39	301
Weighted.............		20	19	53	41	117	44	47	15	21	
Vegetables.............	581										
Unweighted...........		39	29	92	71	167	71	45	26	41	421
Weighted.............		50	37	112	71	165	69	32	23	22	
Fats, oils.............	68										
Unweighted...........		8	4	13	10	17	5	5	1	5	43
Weighted.............		10	7	15	9	14	5	5	2	1	
Sugar, sweets, beverages	315/1										
Unweighted...........		32	15	60	34	76	37	24	18	19	208
Weighted.............		41	18	64	34	84	34	12	14	14	
Total..................	3,702										
Unweighted...........		208	158	639	395	1,053	465	293	210	281	784
Weighted.............		282	201	718	420	1,084	425	261	167	144	

1/ 25 codes discontinued in addition.

Table 3. Use of Not Further Specified (NFS) food codes and number of responses in NFCS 1977-78, 3-day reports only, 48 States

Food group	Total NFS codes used		NFS codes used 1,000 times and over		NFS codes used 1-4 times		Responses for NFS codes	
	Number	Percent[1]	Number	Percent[2]	Number	Percent	Number	Percent
Milk, milk products.......	36	12.9						
Unweighted.............			6	28.6	4	8.7	48,214	25.8
Weighted...............			10	32.3	3	8.1	70,828	27.2
Meat group................	249	23.3						
Unweighted.............			17	53.1	30	11.1	73,306	47.8
Weighted...............			27	62.8	22	10.5	105,644	47.7
Eggs......................	4	8.9						
Unweighted.............			1	20.0	1	14.3	1,943	7.2
Weighted...............			1	20.0	0	0	2,988	7.5
Legumes, nuts, seeds......	5	4.1						
Unweighted.............			1	25.0	0	0	1,941	9.5
Weighted...............			1	20.0	0	0	2,752	9.8
Grain products............	83	9.8						
Unweighted.............			14	26.4	4	2.6	48,299	18.6
Weighted...............			24	31.2	1	1.0	68,812	18.8
Fruits....................	54	14.3						
Unweighted.............			3	21.4	8	8.2	15,375	19.3
Weighted...............			4	20.0	9	10.8	22,047	19.1
Vegetables................	110	18.9						
Unweighted.............			10	25.6	14	12.5	47,806	25.9
Weighted...............			15	30.0	9	11.7	70,579	26.3
Fats, oils................	5	7.4						
Unweighted.............			2	25.0	0	0	30,810	41.7
Weighted...............			2	20.0	0	0	44,505	40.1
Sugar, sweets, beverages..	26	8.3						
Unweighted.............			8	25.0	0	0	52,713	22.3
Weighted...............			12	29.3	0	0	82,143	22.8
Total.....................	572	15.5						
Unweighted.............			62	29.8	61	7.8	320,407	26.2
Weighted...............			96	34.0	44	7.7	470,298	26.6

1/ Percent of total codes in food group. 2/ Percent of codes in food group used 1,000 or more times.

Table 4. Number and percentage of responses for food codes used designated number of times in NFCS 1977-78, 3-day reports only, 48 States

Food group	Total responses[1/]		Number of times food code was used: 1,000 and over		500-999		100-499		Less than 100	
	Number	Percent[2/]	Number	Percent[3/]	Number	Percent	Number	Percent	Number	Percent
Milk and milk products										
Unweighted	186,549	15.3	156,089	83.7	12,383	6.6	13,638	7.3	4,439	2.4
Weighted	260,090	14.7	228,127	87.7	12,255	4.7	15,018	5.8	4,690	1.8
Meat group										
Unweighted	153,515	12.6	76,674	49.9	17,414	11.3	34,495	22.5	24,932	16.2
Weighted	221,620	12.5	136,248	61.5	26,927	12.2	40,139	18.1	18,306	8.3
Eggs										
Unweighted	26,931	2.2	24,644	91.5	531	2.0	1,104	4.1	652	2.4
Weighted	39,594	2.2	36,147	91.3	1,472	3.7	1,137	2.9	838	2.1
Legumes, nuts, seeds										
Unweighted	20,487	1.7	12,002	58.6	2,049	10.0	4,869	23.8	1,567	7.6
Weighted	28,073	1.6	17,015	60.6	3,342	11.9	5,686	20.3	2,030	7.2
Grain products										
Unweighted	259,828	21.3	170,797	65.7	31,248	12.0	41,755	16.1	16,028	6.2
Weighted	365,103	20.7	267,723	73.3	36,250	9.9	48,124	13.2	13,006	3.6
Fruits										
Unweighted	79,844	6.5	53,749	67.3	9,185	11.5	11,097	13.9	5,813	7.3
Weighted	115,206	6.5	84,752	73.6	13,146	11.4	11,122	9.7	6,186	5.4
Vegetables										
Unweighted	184,265	15.1	133,220	72.3	19,569	10.6	21,797	11.8	9,679	5.3
Weighted	268,092	15.2	206,547	77.0	26,610	9.9	25,202	9.4	9,733	3.6
Fats, oils										
Unweighted	73,947	6.1	67,050	90.7	2,832	3.8	2,881	3.9	1,184	1.6
Weighted	109,095	6.2	99,991	91.7	4,996	4.6	3,001	2.8	1,107	1.0
Sugars, sweets, beverages										
Unweighted	236,806	19.4	209,089	88.3	10,768	4.5	12,374	5.2	4,575	1.9
Weighted	359,607	20.4	329,922	91.7	11,728	3.3	13,113	3.6	4,844	1.3
Total										
Unweighted	1,222,172	100.0	903,314	73.9	105,979	8.7	144,010	11.8	68,869	5.6
Weighted	1,766,480	100.0	1,406,472	79.6	136,726	7.7	162,542	9.2	60,740	3.4

1/ Individuals in one-person households weighted in same manner as individuals in larger households.
2/ Percent of total responses from all food groups. Totals may not add to 100 percent because of rounding.
3/ Percent of total responses from each food group. Totals may not add to 100 percent because of rounding.

FROM NUTRIENT DATA TO A DATA BASE FOR A HEALTH AND NUTRITION EXAMINATION SURVEY ORGANIZATION, CODING, AND VALUES--REAL OR IMPUTED

Connie M. Dresser, R.D.
Nutrition Analyst
Division of Health Examination Statistics
National Center for Health Statistics

The National Health and Nutrition Examination Survey (NHANES) and its predecessor, the Health Examination Survey (HBS), collect and utilize data that are obtained by direct physical examination, clinical and laboratory tests, and related measurement procedures. Prevalence data are collected for specifically defined diseases or conditions of ill health; and formative health-related measurement data are collected that show distributions of the total populations with respect to particular parameters such as blood pressure, visual acuity, or serum cholesterol level. Prior to NHANES, the HES programs focused on selected chronic diseases; growth and development and sensory defects of children and youths.

Data on nutrition were first collected in NHANES I and again in NHANES II to monitor changes in nutritional status over time. Data collection methods were selected to record dietary information and a nutrient data base was compiled to process the data. This paper outlines the organization of the original NHANES nutrient data base, its complexities and transitions over the past decade.

ORGANIZATION

The original NHANES dietary component was developed in 1970 by the NHANES nutrition consultant in conjunction with the U.S. Department of Agriculture's Food Composition Research Group. The 24-hour recall method was chosen to record the types and amounts of foods consumed and the vitamin-minerals developments used by the NHANES population on a routine daily basis. The food frequency method was used to depict habitual consumption patterns over the three months prior to their interview. A coding manual and nutrient data base were developed to code and calculate the nutrients contained in the foods reported on the 21-hour recall questionnaires.

After the first year of data collection, the dietary component was evaluated. The format of the nutrient data base was changed to better accommodate the limited headquarters staff and dietary field staff who routinely coded the questionnaires at the collection site. The format was changed to double as a coding manual and the basis for data processing. Numerous problems encountered during the first NHANES were resolved empirically by analysis of the various data collection and processing components.

The nutrient data base consisted of a Model Gram and Nutrient Composition section (Table 1). The model gram section included 1) food group and 2) food code numbers, 3) assigned measure model codes and 4) the food descriptors.

The food groups used in NHANES I were adapted from the Ten-State Survey.

TABLE 1-1 details the groups used to process the 24-hour recall and food frequency forms in both surveys. Foods were assigned to food groups based on descriptive and nutrient content similarities, e.g. milk, cheese, and yogurt are major sources of calcium, phosphorus, and riboflavin. However, in NHANES I, the food items assigned to the 24-hour recall food groups did not necessarily correspond to those foods assigned to the food frequency questionnaire. Discrepancies between the two sets of food groups made it inconvenient to Cross-Check food patterns without first reconstructing food items. Therefore, in NHANES II, the same foods were assigned to the 18 food groups used for both methods.

Food codes available for coding in NHANES I and II, numbered 3630 and 4762 respectively. In NHANES I, the majority of the codes came from USDA food composition handbooks and data tapes, and Tulane University's Dietant Listing, with a minimal number of codes assigned to processed foods from industry or other references (Table 2). Periodically, a recipe was calculated at NHANES headquarters for a mixed dish commonly consumed for which there was no information from USDA. In NHANES II, the number of USDA food codes decreased while the number of processed food codes increased. The decision to add codes was made because food codes provided by USDA, while giving generalized descriptions, did not give sufficient information about packaging, processing, or fortification of brand named foods.

Each food code was assigned a measure model code. This code consisted of 1) an alphabetic symbol used to designate the preferred food model and 2) a gram conversion factor, which is the gram equivalent of one ounce of the corresponding food item. This code was used to convert the ounce equivalents of the food model reported to grams, and then to total nutrients. The alphanumeric description of each food item represented general characteristics, and the method of preparation and packaging, e.g. milk, whole, condensed, canned.

Coding, in NHANES, was conducted at the site of data collection by trained dietary interviewers who held a B.S. in home economics. Their one-week initial training consisted of interviewing techniques, the correct use of questionnaires, food models, the food code manual, and the procedures used to conduct meetings with nutrition professionals in the areas to be surveyed. Followup supervision and interviewer evaluation was conducted by the staff nutrition supervisor.

In both surveys, the Model Gram section of the dietary data base was used as the interviewer coded and verified the majority of food items within 72 hours of collection. Foods left uncoded were identified and coded at headquarters. If the respondent could not recall the exact food consumed or approximate the serving size, no arbitrary substitutes or serving sizes were assigned. Dietary interviewers were responsible for preliminary quality control checks on the collected data. Programmed edits were completed at headquarters to check for missing data or errors in coding, amounts of food or food models reported.

The nutrient composition section of the data base (Table 1) included food group numbers, food code numbers and data on kilocalories, dietary components and nutrients for 100 gram edible portion sizes of respective food codes in the Model Gram section. Data for trace elements and some micronutrients were not added to the data base because the majority of this information was either imputed or available for a limited number of food items.

Updating the NHANES nutrient data base took place twice during both surveys (Table 3). At the beginning of NHANES I (1971), the data sources included:

- food codes from USDA's 1963 data set 8-1. At that time data set 8-1 did not include analyzed nutrient values for total saturated fat, oleic, linoleic acids, and cholesterol. Values for sodium, potassium and phosphorus were limited (code series 0-2700).

- food codes from Tulane University (code series 5000 and 6000).

- unpublished information from USDA which was either imputed data or a recipe calculation (code series 35000).

- nutrient values from USDA's Home and Garden Bulletin #72 and the 1970 edition of Bowes and Church (code series 72000 and 85000).

- a minimal number of nutrient values for foods from IIT Continental Baking Company and other food companies (code series 30000 and 90000).

At the end of NHANES I and before data processing, the nutrient data base was evaluated for 1) programmatic errors, e.g. appropriate decimal placement 2) duplicate codes and nutrient values, 3) missing nutrient values, and 4) appropriate values applicable to the time period. Therefore, in 1975,

- The majority of code series 5000 and 6000 from Tulane University were eliminated because, with few exceptions, the codes were identical to the food codes in the 1963 version of USDA's Handbook #8.

- USDA's 1972 Expansion data tape 8-1-1 of Handbook #8 replaced data set 8-1; however, the adaptation of this tape did not fill in all blanks nor did the literature necessarily provide additional needed information.

- all other remaining code series were updated, where possible, with nutrient values from appropriate sources. On the NHANES I Model Gram/Nutrient Composition data release tape, a symbol of 8888888 was used to indicate that the data were unavailable.

At the beginning of NHANES II (1976), the following changes were made:

- all food codes and nutrient values on the USDA data tape 8-1-1 were replaced with food codes and values from USDA's Handbook #456. The nutrient values were taken from data tapes 456-1 and 456-2 (code series 0-2700).

- remaining Tulane University codes were updated with values from USDA's Home and Garden Bulletin #72, the 12th edition of Bowes and Church, USDA's Handbook #456 and the National Heart Study (code series 5000 and 6000).

- in 1978, NHANES II went to Hawaii and the nutrient values for foods reported were taken from research bulletins and circulars applicable to foods grown on the island (code series 8000).

- NHANES I foods taken from the IIT Continental Baking Company were updated with information from the company or with nutrient values of similar products used in USDA's 1977/78 Individual National Household Food Consumption Survey (code series 30000).

- USDA baby food codes were replaced with nutrient values from Baker/Beechnut and Gerber baby food companies (code series 10000 and 41000).

- USDA's unpublished information was updated with calculated values or recipes (code series 35000).

- foods taken from USDA's Home and Garden Bulletin #72 and Bowes and Church were updated with current editions of both publications (code series 72000 and 85000).

- new commercial foods added to the data base were assigned codes between 90000 and 9S000.

At the close of NHANES II, only values for foods reported twenty or more times were updated to reflect knowledge of the nutrient content of foods between 1976 and 1980. The update was limited due to staff and time constraints. If values from industry were unavailable, values were substituted for similar foods in USDA's Handbook #456, and revised editions of Handbook #8, sections, 1-6. If no appropriate values could be found, a symbol of 9999999, not zero, was assigned on the Nutrient Composition data release tape to indicate a thorough search had been made but no reliable information was available.

NUTRIENT VALUES--REAL OR IMPUTED?

Undating the NHANES dietary data base, in both surveys, was hampered by the unavailability of nutrient values applicable to the time periods. This is important to remember when deciding whether the values addressed are "real or imputed". For instance, years ago most of the nutrient data published by the Nutrient Composition Laboratories of the Department of Agriculture were based on direct chemical analyses. The majority of this data has been replaced with composite or composition values from industry, independent laboratories, and scientific or technical literature. Although the nutrient values on USDA data tapes and in handbooks are suitable for determining intakes on kinds and amounts of foods consumed and reported in today's dietary surveys, the composition values represent foods as used throughout the country on a year-round basis. Values appropriate at one time may not apply during a latter period when new or different varieties predominate; when production, manufacturing or preparation processes have changed, or when different ingredients or improved methods of analysis are used. However, while numerous food companies have been most cooperative with the dissemination of their product analyses, users of this data are unaware of methods of analysis or frequency of imputation. It is acknowledged by food companies and most data base developers that the label or fact sheet data provided to the consumer are often based on the percent of the U.S. RDA's or on the NRC RDA standards. When label data are compared to corresponding 100 gram edible portion data, there are discrepancies. Both of these data sources offered limited data on sugars, fiber and nutrients such as zinc, copper, magnesium, folacin, vitamin B6, vitamin B12, and pantothenic acid. Concurrently, the effects of processing and packaging methods and product fortification make it impossible to define the "real values" of foods used in large-scale studies. These unavoidable conditions make it important to know the limitations and caveats of nutrient data bases when analyzing dietary data.

Another important issue to address is the effect of missing nutrient values in the data base. For example, Table 4 suggests that the multitude of missing values accompanying the brand named foods in both the NHANES surveys should have a negative effect on the total nutrient intakes. Before assuming this is true, the data tape user should know how often the foods with missing values were consumed. Table 4 also suggests that the data from sources other than food manufacturers provided more information on nutrient values in foods in NHANES II than in NHANES I. However, without further investigation, the data tape user does not know the quality of this additional data. For example, in NHANES I, 1972 USDA data set 8-1 gave no options for coding foods cooked with salt versus no salt or foods prepared with butter, oil or margarine. This information was available in data set 456-3 during NHANES II. These changes in the data base, in part, contribute to the observed changes in nutrient intakes between the NHANES surveys.

Because data tape users or developers rarely know what percentage of the data are imputed or "real", careful monitoring of the data base is essential. Changes made in the data base will partially explain the shifts in nutrient intakes by population sub-growth in surveys. In NHANES, changes in nutrient intakes between surveys can be attributed to the differences in frequency of food item consumption, total grams of food consumed, changes in food fortification, and to the changes made in the nutrient data base.

It is important to remember that NHANES is not in the "nutrient data base business". Our data base is used primarily by researchers who are addressing questions related to health, nutrition and disease. However, as a data base developer, we must be knowledgeable of the limitations of the nutrient values which will eventually be used by researchers studying the interrelationships of diet and health. Being aware of these caveats is not enough if minimal progress is made to improve the state of the art--a problem which has long faced large-scale surveys. Dietary methodologies and nutrient data bases currently used are still too costly and time consuming. However, work is being done to improve the dietary collection methods and methods of analyzing nutrients. One methodological example is the "core frequency questionnaire" used in the NHANES I Epidemiological Followup Survey at the National Center for Health Statistics. The format includes 100 precoded food items most often consumed in both of the NHANES surveys. The selection of the food items was based on the number of times the food item was reported, total gram consumption and its contribution to the majority of nutrient intake. To calculate median nutrient values for daily and weekly consumption of these foods, a formula using frequency count, reported serving size information and an aggregated nutrient values per food item is necessary. This procedure is relatively simple until nutrient values are addressed. The nutrient data base needed to process this data should be standardized. Standardization of the base is an issue that needs to be evaluated and set into motion before substantial progress can be made in linking dietary intake data to health and disease.

References

1. Baker-Beechnut Baby Food Company, CanoJoharie, New York. 9/1975.

2. Bowes and Church; Food Values of Portions Commonly Used. 1970. 11th edition.

3. Bowes and Church;: Food Values of Portions Commonly Used. 1975. 12th edition.

4. Centers for Disease Control: Ten-State Nutrition Survey in the United States, 1968-1970. Part V. Dietary. DHES Pub. No. (HMS) 72-8133. Atlanta, Georgia.

5. Contemporary Nutrition, Watt, B.K.: Tables of Food Composition: uses and limitations. Vol. 5, No. 2. Feb. 1980.

6. Department of Health, Education and Welfare, Leung, W-T, Butrum R., Roa, M. and Polacchi, W.: Food Composition table for use in East Asia. DHEW Pub. No. (NIH) 79-165. Dec. 1972 - reprint 1978.

7. Gerber Baby Food Company, Fremont, Michigan. 9/1975.

8. ITT Continental Baking Company, P.O. Box 751, Rye, New York. 1970-71.

9. National Center for Health Statistics, Pearce, N.; Data Systems of the National Center for Health Statistics. <u>Vital and Health Statistics</u>. Series 1, No. 16. DHHS Pub. No. (PHS) 82-1318. Public Health Service. Hyattsville, Maryland, December 1981.

10. National Center for Health Statistics: Public users tape documentation: NHANES I - Model Gram/Nutrient Composition Tape Cat. No. 1702, NTIS No. PB-296027. Public Health Service. Hyattsville, Maryland, May 1981.

11. National Center for Health Statistics: Public users tape documentation: NHANES II - Model Gram/Nutrient Composition Tape, Cat. No. 5704, NTIS No. PB-82-142613. Public Health Service. Hyattsville, Maryland, Feb. 1982.

12. National Institutes of Health: The National Heart Study. 1965.

13. Tulane University; Tulane Dietant Listing. 1969.

14. University of Hawaii, Agriculture Experiment Station, Wenkam, N. and Miller, C.; Bulletin 135 - C85 - Composition of Hawaii Fruits. Reprint October 1973.

15. University of Hawaii, Agricultural Experience Station, Standal, R., Bassett, D., Policar, P., and Thom, M.; Bulletin 146 - Fatty Acids, Cholesterol, and proximate composition of certain prepared and unprepared foods in Hawaii. May 1975.

16. University of Hawaii, Agricultural Experiment Station, Miller, C. and Branthoover, B.: Circular 52 - Nutritive values of some Hawaii foods. Reprint Dec. 1963.

17. U.S. Department of Agriculture, Agriculture Research Service, Consumer and Food Economics Research: Composition of foods--raw, processed, prepared. Handbook No. 8 revised. Data set 8-1. Dec. 1963.

18. U.S. Department of Agriculture, Agriculture Research Service, Consumer and Food Economics Research; Composition of foods--raw, processed, prepared. Expansion No. 1 of data published in Agriculture Handbook No. 8. Data sets 8-1-1, 8-2-1. April 1972.

19. U.S. Department of Agriculture, Agriculture Research Service, Consumer and Food Economics Research: Expansion of data published in Nutritive Values of American Foods in Common Units. Handbook No. 456. Data sets 456-1, 456-2. May 1977.

20. U.S. Department of Agriculture, Agriculture Research Service, Consumer and Food Economics Research: Composition of foods. Handbook Nos. 8-1 through 8-6, Nov. 1976-Feb. 1980.

21. U.S. Department of Agriculture, Agriculture Research Service: Nutritive Values of foods. Bulletin 72. 1970.

22. U.S. Department of Agriculture, Agriculture Research Service, Science and Education Administration: Nutritive Values of foods. Bulletin 72, 1977.

23. U.S. Department of Agriculture, Agriculture Research Service Consumer and Food Economics Research: 1977/78 Nationwide Food Consumption Survey on Individuals. Tape PB82138504, released 1-1980.

24. U.S. Department of Agriculture, Agriculture Research Service, Consumer and Food Economic Research, Food Composition Research Group: unpublished data. 1971.

TABLE 1

NHANES NUTRIENT DATA BASE

- MODEL GRAM TAPE
  - FORMAT
    - Food Group Number
    - Food Code Number
    - Measure Model Code
    - Description of Food
  - FOOD GROUPS - 24-HOUR RECALL
    - NHANES I
      1. Milk & Milk Products
      2. Meat
      3. Poultry
      4. Organ Meats
      5. Fish and Shellfish
      6. Eggs
      7. Soups, Gravies and Sauces
      8. Fats and Oils
      9. Legumes and Nuts
      10. Cereals and Grain Products
      11. Fruits and Vegetables
      12. Sugar and Foods Primarily Sugar
      13. Desserts, incl. ice cream
      14. Miscellaneous
      15. Mixed Protein Dishes
    - NHANES II
      1. Milk & Milk Products
      2. Meat
      3. Poultry
      4. Organ Meats
      5. Fish and Shellfish
      6. Eggs
      7. Soups, Gravies and Sauces
      8. Fats and Oils
      9. Legumes and Nuts
      10. Cereals and Grain Products
      11. Fruits and Vegetables
      12. Sugar and Foods Primarily Sugar
      13. Desserts and Sweets
      14. Miscellaneous
      15. Mixed Protein Dishes
      16. Alcoholic Beverages
      17. Sugar Free and Low Calorie Beverages
      18. Salty Snacks
- NUTRIENT COMPOSITION TAPE
  - FORMAT
    - Food Group Number
    - Food Code Number
    - Kilocalories
    - Dietary Components
      - CARBOHYDRATES
      - PROTEIN
      - FAT
      - *TOTAL SATURATED FATTY ACIDS
      - *OLEIC ACID
      - *Linoleic Acid
      - *Cholesterol
    - Nutrients
      - Calcium
      - Phosphorus
      - Sodium
      - Potassium
      - Iron
      - Vitamin A
      - Vitamin C
      - Thiamin
      - Riboflavin
      - Preformed Niacin

*NHANES II NUTRIENT COMPOSITION TAPE ONLY

TABLE 1-1

NHANES FOOD GROUPS

NHANES I (1971-1974)		NHANES II (1976-1980)
24-Hour Recall Food Groups	**Food Frequency Food Groups**	**24-Hour Recall/Food Frequency Food Groups**
1. Milk and Milk Products	1. Milk	1. Milk and Milk Products
	Whole	Whole Milk
	Skim	Skim Milk
2. Meats	2. Meat and Poultry	Ice Cream
3. Poultry	3. Fish and Shellfish	Cheese and Cheese Dishes
4. Organ Meats	4. Eggs	2. Meats
5. Fish and Shellfish	5. Cheese and Cheese Dishes	3. Poultry
6. Eggs	6. Dried Beans and Peas	4. Organ Meats
7. Soups, Gravies and Sauces	7. Fruits and Vegetables	5. Fish and Shellfish
	those rich in Vitamin A	6. Eggs
	those rich in Vitamin C	7. Soups, Gravies and Sauces
8. Fats and Oils	8. Bread and Cereals	8. Fats and Oils
9. Legumes and Nuts	9. Butter and Margarine	9. Legumes and Nuts
10. Cereal and Grain Products	10. Desserts and Sweets	10. Cereals
		Grain Products
11. Fruits and Vegetables	11. Candy	11. Fruits and Vegetables
12. Sugar and Foods Primarily	12. Beverages	those rich in Vitamin A
sugar		those rich in Vitamin C
13. Desserts, inc. ice cream	13. Snack Foods	12. Sugar and Primarily Sugar Products
14. Miscellaneous		13. Desserts and Sweets
15. Mixed Protein Dishes		14. Miscellaneous
		15. Mixed Protein Dishes
		16. Alcoholic Beverages
		Beer
		Wine
		Distilled liquor
		17. Sugar Free and low Calorie Beverages
		Artificially sweetened drinks
		Coffee or Tea
		18. Salty Snacks

TABLE 2

SOURCES OF CODES REPORTED IN NHANES BY CODE SERIES

CODE SERIES	USDA HB.# #8 456 0-2700		Tulane Univ. 5000/6000		Hawaii & East Asian 8000		Continental Baking Co. 30000-31000		USDA Confid. information 35000		Baker/ Beechnut 40000	Gerber's 41000	USDA Bull. #72 72000		Bowes&Church 85000		Commercial 90000-95000	
FOOD GROUPS NHANES	I	II	I	II	I	II	I	II	I	II	II	II	I	II	I	II	I	II
1. DAIRY	51	50	7	7	-	0	0	0	0	0	2	2	1	1	2	2	5	81
2. MEAT	134	165	22	6	-	0	0	0	1	0	3	9	10	0	0	0	0	31
3. POULTRY	39	43	3	2	-	0	0	0	7	0	3	5	0	0	0	0	0	5
4. ORGAN MEATS	19	23	1	5	-	0	0	0	0	0	0	1	0	0	0	0	0	2
5. FISH & SHELLFISH	81	87	7	5	-	11	0	0	0	0	0	0	0	0	0	0	0	3
6. EGGS	9	11	0	1	-	0	0	0	0	0	0	0	0	0	0	0	0	3
7. SOUPS & SAUCES	66	58	10	9	-	1	0	0	0	0	0	0	0	0	0	0	0	56
8. FATS & OILS	37	34	6	7	-	0	0	0	1	1	0	0	7	7	0	0	0	41
9. LEGUMES/NUTS/SEEDS	53	51	2	4	-	0	0	0	0	0	0	0	0	0	1	1	4	25
10. CEREALS & GRAIN PRODUCTS	191	128	44	23	-	2	1	2	7	6	6	18	2	2	0	0	55	174
11. FRUITS & VEGETABLES	432	447	41	25	-	17	0	0	5	5	43	48	0	0	0	0	8	53
12. SUGAR & SUGAR PRODUCTS	76	77	11	21	-	0	0	0	0	0	0	0	0	0	0	0	2	15
13. DESSERTS	98	84	19	25	-	0	14	0	13	10	6	15	1	1	0	0	37	154
14. MISCELLANEOUS	40	29	7	1	-	0	0	0	2	0	0	0	2	1	2	0	4	26
15. MIXED DISHES	58	37	17	10	-	1	4	0	4	3	14	38	0	0	0	0	6	117
16. ALCOHOLIC BEVERAGES	7	8	1	3	-	0	0	0	0	0	0	0	0	0	0	0	0	0
17. LOW CALORIC BEVERAGES	11	11	1	0	-	0	0	0	2	3	0	0	0	0	0	0	5	21
18. SALTY SNACKS	4	4	3	3	-	0	2	2	0	0	0	1	0	0	0	0	0	0
TOTALS	1406	1347	202	157	-	32	21	4	42	28	77	137	23	12	5	3	126	807

Survey Totals: NHANES I - Codes available for coding: 3630, reported 1825 = 50%
NHANES II - Codes available for coding: 4762, reported 2604 = 55%

TABLE 3

NHANES NUTRIENT DATA BASE UPDATE

CODE SERIES	NHANES I (1971)	NHANES I(1975)	NHANES II(1976)	NHANES II(1980)
0-02700 USDA's Handbook #8	Data set:8-1, 1963 version	1972 Expansion of Handbook #8, Data set:8-1-1	USDA's Handbook #456-3 Data set: 456-1 and 456-2	
5000/6000 Tulane University's Dietant Listing	Data from 1947-1969 majority of codes were duplicates of USDA's Hb. #8.	Duplicate codes were eliminated.	5000/6000 codes from NHANES I were updated with: with: o USDA's Home & Garden Bulletin #72, 1971 edition o Bowes & Church, 12th ed. o USDA's Handbook #456 o The National Heart Study	Values updated with same sources, where appli-, cable, also the Nationwide Food Consumption Survey data base developed for individual intakes. (HFCS)
08000 Hawaiian Foods and East Asian Foods	**	**	o Research Bull. #135 - Composition of Hawaii Fruit o Research Bull. #146 - Fatty Acids, cholesterol, and proximate composition of certain prepared and unprepared foods in Hawaii o Circular #52 - Nutritive values of some Hawaiian Foods o Food composition table for use in East Asia	
30000 ITT Continental Baking	1971 values from company	update from company	HFCS and commercial data	
35000 USDA Confidential Information	1971 values from USDA's Food Composition Research Group	update from USDA	update from USDA	HFCS data and 456-3

TABLE 3 - continued

NHANES NUTRIENT DATA BASE UPDATE

CODE SERIES	NHANES I(1971)	NHANES I(1974)	NHANES II(1976)	NHANES II(1980)
40000 Baker/Beechnut Baby Food Company	**	**	1975 data from company	update from company
41000 Gerber Baby Food Company	**	**	1975 update from company	update from company
72000 USDA's Home & Garden Bulletin #72	1970 edition	updated with the 1971 slightly revised edition	1971 edition	1977 edition
85000 Bowes & Church	1970 edition(11th)	1975 edition(12th)	1975 edition	
90000 Commercial products	data from companies	update from companies	90000-95000 commercial products	o update from appropriate companies o USDA's Hb. #456-3 o HFCS data base o USDA's Hb. #8, sections 1-6

** Code series not included in survey.

TABLE 4

NUMBER AND PERCENT OF FOODS REPORTED IN NHANES WITH MISSING NUTRIENT VALUES

	NHANES I				NHANES II			
NUTRIENT	**BRANDS*(9)**		**OTHERS**		**BRANDS*(21)**		**OTHERS**	
	n=147	%	n=1678	%	n=1025	%	n=1579	%
CALORIES	1	0.7	6	0.4	0	0.0	0	0.0
CARBOHYDRATES	3	2.0	19	1.1	6	0.6	1	0.1
PROTEIN	12	8.1	18	1.1	6	0.6	3	0.2
FAT	14	9.5	21	1.3	6	0.6	4	0.2
SATURATED FATTY ACIDS	**	**	**	**	627	61.2	70	4.4
OLEIC	**	**	**	**	643	62.7	70	4.4
LINOLEIC	**	**	**	**	639	62.3	72	4.6
CHOLESTEROL	**	**	**	**	552	53.9	82	5.2
CALCIUM	19	12.9	105	6.3	32	3.1	16	1.0
PHOSPHORUS	20	13.6	145	8.6	267	26.1	31	2.0
SODIUM	18	12.2	269	16.0	184	18.0	67	4.2
POTASSIUM	52	35.4	296	17.6	288	28.1	65	4.1
VITAMIN C	77	52.4	484	28.8	144	14.1	35	2.2
VITAMIN A	74	50.3	285	17.0	117	11.4	32	2.0
THIAMIN	15	10.2	149	8.9	67	6.5	22	1.4
RIBOFLAVIN	13	8.8	154	9.2	57	5.6	19	1.2
PREFORMED NIACIN	19	12.9	164	9.8	74	7.2	24	1.5
IRON	18	12.2	123	7.3	42	4.1	20	1.3
	Total: n = 1825				Total: n = 2604			

* processed foods with known brand names
** *nutrients not available for the majority of foods*

SOURCES OF VARIATION IN NUTRIENT COMPOSITION DATA

Joanne M. Holden
Nutritionist
Nutrient Composition Laboratory
U.S. Department of Agriculture

As consumers and professionals, we know that the American marketplace offers thousands upon thousands of food products. Products in the retail market may be commodity items--fresh produce, fresh meats, coffee beans--or they may be brand-name, complex processed recipe formulations. Also, within each product category there are many items and brands from which to choose. For example, the category salty snack foods includes cheese twists, pretzels, potato chips, and crackers of all kinds. The flavors, shapes, and textures of these are limited only by the imagination of product development teams. Each product may have from 1 to 30 different ingredients. The various ingredients may have been extracted, dehydrated, enriched, etc.; treated with a list of chemicals inherent in these processes; then fried or baked; packaged and shipped from coast to coast.

This cornucopia of products may be exciting to the consumer, but it represents a giant-size headache for those of us who sample the food system in an attempt to prepare representative samples of food products for subsequent nutrient analyses. As analysts and contractors, we cannot afford the time and money to determine the nutrient content of every product (nor would we want to). As nutritionists or dieticians, we want to do as little coding as possible. If the nutrient content of potato chips varies little from brand to brand, then we don't need to code each brand of potato chip. How can we, as professionals who are concerned about the nutrient composition of foods, determine which products are the same and which are different?

I would like to share some of my observations and experiences concerning the evaluation of variations in nutrient composition data and the implications of this variation for food sampling.

For the purposes of this discussion, variation can be defined as deviation from a central position or value. If the amount of deviation is small, one can have higher confidence in the central value than if the deviation is relatively large. The ultimate goal in the area of nutrient composition of foods is to have statistically meaningful data for all nutrients in all foods. However, in the real world of limited budgets and deadlines, we must develop a specific strategy, determined by priorities, for sampling the foods which people eat. The Nutrient Composition Laboratory has developed just such a strategy. The strategy includes a systematic approach to determining which foods and which nutrients should be analyzed (G. R. Beecher and J. T. Vanderslice).

Using the strategy we have developed for selecting food products and nutrients, we have sampled beef and pork, fruit juices, yogurt, salty snack foods, and cereals. Planning a nutrient composition project for a selected product type is an iterative process. One must define the potential sources of variation for that product in order to evaluate the extent of variability. Then, one can incorporate those data into the next phase of the project to estimate mean values for nutrients.

Table 1 shows potential sources of variation in all aspects of the sampling process. Some of these sources can be controlled to limit their effects on total variation; others are peculiar to the product and food supply and must be left to vary as they will. However, important but uncontrolled sources of variation should be quantifiable.

TABLE 1: POTENTIAL SOURCES OF VARIATION RELATIVE TO NUTRIENT COMPOSITION DATA

- PRODUCT CHARACTERISTICS
- PLANT LOCATION
- QUALITY ASSURANCE PROGRAMS
- SEASONAL EFFECTS
- DISTRIBUTION TECHNIQUES
- SAMPLE SELECTION
- SAMPLE PREPARATION
- ANALYTICAL METHOD

Product characteristics which define the generic product are of central importance (Table 2).

TABLE 2: PRODUCT CHARACTERISTICS

- PRODUCT TYPE
- BRAND NAME
- FOOD SOURCE
- PART OF PLANT OR ANIMAL
- PHYSICAL STATE, SHAPE, OR FORM
- PRESERVATION METHOD
- TREATMENTS APPLIED TO PRODUCT
- USER GROUP
- PACKAGING

"Product Type" refers to the food group to which the product belongs. Products belonging to the same group may have common consumption, functional, or manufacturing characteristics. A product type may include specific products or brands which are different from each other in other characteristics. These differences or multiple values can affect the extent of variation in some nutrients. Certain types of products which can be differentiated primarily by brand can differ due to variations in formulation and ingredients. For others the source of the primary ingredient may vary. For example, in a study of ready-to-eat (RTE) cereals the main ingredients of various flaked cereals may be corn, wheat, or oats.

"Food Source" denotes the individual plant, animal, or chemical from which the food product or its major ingredient is derived. Food source also includes the natural variability attributable to species, variety, and cultivar. If we were dealing with meat products, we might need to define the various products in terms of the species used in the preparation of these products. Within a product class, such as bacon, there is great variation in the ratio of physical components, fat and lean. When we sampled bacon, we compensated for the large amount of variability by compositing several pounds of each brand, instead of selecting one package. Our purchasing agents were strongly advised to select those several packages at random from the meat case instead of selecting the leanest packages.

"Part of Plant or Animal" refers to the anatomical part of the plant or animal source used to prepare the product. For yogurt, the part of animal would be milk; for salty snacks, the part may be endosperm. For certain meat products, the definition of cuts used would be important. As I mentioned earlier, factors for a given product class become more important when they have multiple values.

"Physical State, Shape, or Form" of the product includes particle size, product shape, and viscosity. Do pretzel sticks have more salt on the outside than the large, hard Dutch pretzels?

"Degree of Preparation" denotes the degree of cooking and preprocessing, factors which may affect the stability of nutrients within a product. The method of cooking for prepared food products is also included. The effect of various cooking methods on nutrient retention in foods can influence the extent of variability in nutrients inherent in the product. In addition, various cooking methods can contribute significant amounts of ingredients which are nutrients. For example, cheese curls or twists may be fried or baked. One would expect the two products to differ in their fat content. This may or may not be true; analysis of the two products would provide the answer. Across the Product Type Salty Snack Foods various methods of preparation and cooking will contribute to the nutrient variability of the product class.

"Treatments Applied . . ." include enrichment, as well as the application of specific processes such as fermentation and the addition of secondary ingredients. Secondary ingredients include sugar, fat, salt. Certainly variation in the amount and source of these ingredients affects the nutrient composition of the product.

"Preservation Method" refers to the primary method used to prevent or retard spoilage. What are the effects of the preservation method upon the nutrient composition of foods? In a study of fruit juices, we found significant differences in the vitamin C content of chilled reconstituted juices and frozen concentrates.

The definition of specific product characteristics is only part of the evaluation process. We must also define other aspects of the manufacturing and distribution process. The location of manufacturing plants for classes and brands of products can be an important source of variation. Certain brands of cookies are manufactured in one plant for the whole country while other brands of cookies are produced at several plants. Other products, such as dairy products, sausages, and bacon are produced in many regional plants all over the U.S. Many of these products are not distributed nationwide.

We found that some brands of bacon are produced at various plants within a meat processing company. In many cases the brine specifications and pump rates for that brand are the same from plant to plant. However, at least one brand was marketed nationwide and produced by different meat processors. All processors producing that brand were not using the same brine specifications and pump rate. However, we could identify only the specific brand-plant combination that we picked up; fortunately we had access to the approved brine specifications for that product. Since bacon is so complex, we limited our objective to the estimation of a range of concentration for selected nutrients. The selected nutrients will be determined in products manufactured according to different brine specifications and pump rates. So, you see, the aspect of manufacturing plant location can be quite complex.

Quality assurance programs within each plant affect the amount of variability in the nutrient composition of a product. Fortunately, economic considerations force portion control and product standardization and, in turn, reduce product variability within each plant. A measure of lot-to-lot variability would be an indicator of quality control.

At this point, it is important to emphasize that we must consider the amount of variation, nutrient by nutrient, or at least define an indicator nutrient for that product type. As you know, some nutrients are affected more by a lack of standardization than others. Also, if the nutrient is inherent or bound to the raw ingredient, there may be less cause for concern than if the nutrient is added by spraying or mixing a solution or tablet into the recipe. We have observed large differences between labeled iron content and the analyzed iron content of several brands of RTE cereals. Let me emphasize that analytical values for most brands that we sampled were within ±20% of the label value. However, for certain brands the iron values were between 160% and 260% of the label value. Mean values were derived from analysis of one package from each of three lots. Coefficients of variation for the mean values varied from brand to brand. In our experience, we have found that lot-to-lot variation or within-plant variation is of little concern except for special studies of quality assurance.

Moving from quality assurance, we may discuss the effects of season on nutrient variation. In the Nutrient Composition Laboratory we have not dealt with seasonal variation to any great extent. Many foods are harvested at one time and then stored until they are used in the manufacture of products. Other foods are harvested and are partially or fully processed soon after harvest. Still others have long shelf stability and are not affected by season. For certain projects, seasonal effects are worthy of consideration.

Having defined the characteristics of the product class to be sampled, and other relevant factors, one must determine the sampling plan to insure selection of representative samples. The use of statistically based sampling techniques insures the selection of samples which are representative of the population of products to be evaluated. Careful planning to determine the types and numbers of samples to be selected can prevent the collection of a biased sample and subsequent erroneous estimates of the extent of variation within the population of products.

After representative samples have been selected they must be prepared for nutrient analysis. Preparation may include dissection, cooking, weighing and other specific processing techniques. These procedures are usually followed by homogenization according to predetermined and carefully tested protocol under controlled conditions (time, temperature, and humidity). Subsamples of the ground food product are stored in appropriate containers according to conditions prescribed by the specific nutrient methods.

The final source of variation is nutrient analysis including all steps of the analytical procedure. The use of standard reference materials and control samples to facilitate accuracy, precision, and specificity of the technique must be considered (W. Wolf; J. M. Harnly et al.). Careful handling of the samples at every step with attention to timing, reagents, instrumentation, and environmental conditions will minimize the spurious and random variation which can occur.

These procedures for sample preparation and nutrient analysis should be evaluated and standardized to control unwanted variability which can mask the central issue, "What are the sources and extent of variation in nutrient composition data which are attributable to the food product, from the point of manufacture through distribution to the point of consumption?"

The evaluation of the sources and the extent of variation in nutrient composition data can permit the confident estimation of mean values of nutrients in foods. In addition, knowledge of the sources and amount of variation permits the maximum use of financial and human resources.

This general discussion of the sources and extent of variation in nutrient composition data will be discussed in greater detail in future papers to be published in the Journal of Food Science. References which have been included should supplement the present discussion and clarify certain general points.

Literature Cited

Beecher, G. R. and J. T. Vanderslice, 1983. Determination of Nutrients in Foods. In K. K. Stewart, Ed., "Modern Methods of Food Analysis." Seventh IFT-IUFOST BASIC Symposium held at 43rd Annual Meeting of International Food Technology, New Orleans, LA.

Harnly, J. M., W. R. Wolf, and N. J. Miller-Ihli, 1983. Quality Assurance of Analysis of Inorganic Nutrients in Foods. In K. K. Stewart, Ed., "Modern Methods of Food Analysis." Seventh IFT-IUFOST BASIC Symposium held at 43rd Annual Meeting of International Food Technology, New Orleans, LA.

Wolf, W. R., Ed. "Biological Reference Materials: Availability, Use, Needs for Nutrient Analysis," John Wiley and Sons, NY, 1984 (In Press).

A MICROCOMPUTER-BASED INTERACTIVE MODEL FOR COLLECTION AND CODING OF DIETARY DATA

I. Marilyn Buzzard, Nutrition Coding Center, Division of Biometry
Lynda B. M. Ellis, Division of Health Computer Sciences
Elaine D. H. Asp, Department of Food Science and Nutrition
University of Minnesota

INTRODUCTION

Recent technological advances in computer hardware and software have made possible the automation of procedures for collecting and coding dietary data. A model for microcomputer-based interactive programming is being developed at the University of Minnesota Nutrition Coding Center (NCC) to improve the accuracy and efficiency of dietary data collection and coding procedures.

A number of other investigators have successfully implemented interactive systems for dietary data collection and analysis (Witschi et al. 1981; Evans & Gormican, 1973; Slack et al. 1976). However, none of these systems documents dietary intake at the level of precision required by studies using the present NCC system for investigation of diet-disease relationships.

This paper will begin with a brief description of the present NCC system and will then proceed to describe the initial stages of development of the model for computerizing the coding component of the NCC dietary analysis system.

THE PRESENT NCC SYSTEM

The NCC system was developed in 1974 under NHLBI contract to provide standardized procedures for collecting, coding, and calculating nutrients for 24-hour dietary recalls for two collaborative nationwide cardiovascular studies, the Multiple Risk Factor Intervention Trial (MRFIT) and the Lipid Research Clinics (LRC) programs. Standardization of collection and analysis procedures for the 34 clinical centers involved was achieved through centralized training of nutritionist interviewers and centralized coding at the NCC.

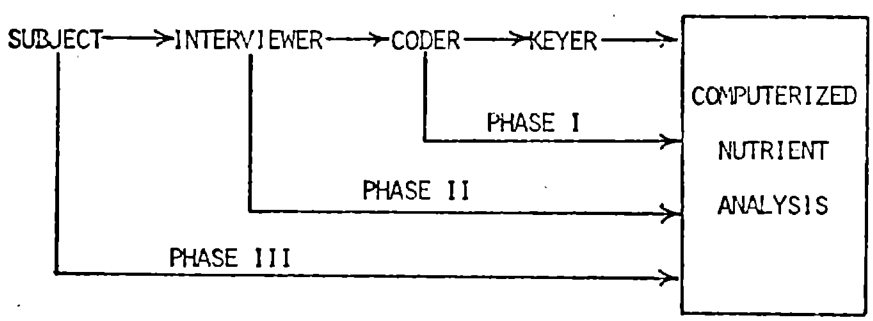
Detailed descriptions of the NCC system have been presented by Dennis et al. (1980) and Tillotson et al. (1981). Briefly, the system provides for detailed documentation of food intake through the use of a 500-page coding manual which includes several hundred coding rules to provide consistency in coding when amounts, preparation methods or food characteristics are documented as unknown (such as amount of creamer used in coffee or unknown cut of beef). The system also includes a Recipe File of over 550 recipes, a Manufacturers' File containing nutrient and ingredient information on thousands of commercial products, and an "Uncodables" File for standardization of infrequent coding decisions. The codebook and nutrient data base are continually updated to reflect changes in the marketplace. Designed specifically to characterize the quality and quantity of dietary fat and cholesterol, the system allows detailed specification of the type and brand of fat used in recipes, fats added in food preparation, and estimates of amount of fat absorbed in various food preparation methods. A comprehensive system of internal and external quality controls has been incorporated into the coding and analysis procedures to ensure accuracy and consistency over time.

In 1977, the NCC system was made available to other studies, and modifications have been made to the individual research needs such as expansion of the system to include processing of multiple-day diet records, diet histories, and quantified food frequencies. In addition to cardiovascular studies, various studies involving cancer, diabetes, hypertension, gastroenterology and other diet-related diseases have used the NCC system since 1977. Specialized coding to accommodate the special research needs of each new study, such as quantification of sodium or fiber intake, have led to ever increasing numbers of new coding rules and the expansion of the nutrient data base. As the system grows in complexity, it has become increasingly difficult and costly to maintain the present level of accuracy and standardization. An automated system would free the food coder from the cumbersome manual coding procedures while maintaining the meticulous attention to coding detail. Thus, a model is being developed for a microcomputer-based interactive system that will incorporate all of the documentation and coding features of the present NCC system, including procedures for verification, error checking, and other quality controls.

MODEL FOR COMPUTERIZATION OF THE NCC SYSTEM

NCC staff members are collaborating with the University of Minnesota Health Computer Sciences Division and the Department of Food Science and Nutrition in developing the computerized model. The Computerized NCC System (CNCC) model involves three phases of development as indicated schematically in Figure 1 below:

Figure 1. Schematic diagram of the steps involved in dietary data collection and analysis. The arrows for phases I, II, and III indicate the successive steps to be eliminated in the development of the microcomputer-based model.



Phase I, the coder-oriented system, will include the development and evaluation of an interactive coding system that will combine coding and keying into a single-step process. Phase II, the interviewer-oriented system, will combine interactive coding with interviewing, allowing improved documentation of food intake while eliminating the need for a food coder. Phase III, the subject-oriented system, will involve modification of the interactive system for use by untrained subjects, thus eliminating the need for the interviewer, as well as for the coder and keyer.

This paper is limited to the development of Phase I, the coder-oriented component of the system. Most of the software to be developed in Phase I will be utilized, with appropriate modifications, in the succeeding phases. Testing at the end of Phase I will allow comparison of the CNCC system with the established NCC system.

Phase I development will include interactive programming that will incorporate the procedures presently used by the NCC food coders. Foods will be identified through an interactive menu-driven branching system based on food acronyms arranged in subsets of food grouping hierarchies. Some 100 to 200 major foods and food groups, designated by a one or two letter acronym (e.g. BF = beef; EG = egg; VG = vegetable; C = coffee; T = tea) will constitute the uppermost level of the hierarchy. Each group will be associated with a menu of its subgroups (either type or form of food or brand name). Each subgroup will be similarly structured using increasingly more specific descriptions and ending with the most precise identification of the food item including method of preparation, fats and salt used in preparation, and fat and salt added at the table.

An illustrative example of interactive coding using such a system is shown in Figures 2-7. Figure 2 is an example of documentation provided on a typical dietary record form. In Figure 3, the item salmon, pink, fresh and be identified either by three successive menu selections or by entering FS.SAL.PK.FR in a single sequence. Figure 4 indicates the amount of fish consumed either by weight, by known volume, or by dimensions for calculating the volume (such as length, width, and thickness of a rectangular serving). Preparation methods are indicated in Figures 5-7. Selection of BST, the present NCC "prep code" for oven cooked and basted, prompts a menu for type of fat used in the preparation as shown in Figure 5. If the fat used is unknown, documentation as to whether the food was consumed at home or in a commercial establishment is used to aid in selecting an appropriate fat. Margarine is an item that must be specified by brand and type, if known, as illustrated in Figure 6. The "Margarine Guide" for the present NCC system indicates which of the 44 margarine entries in the Food Table should be used for each of the several hundred brand name margarines based on type and polyunsaturated to saturated fatty acid (P/S) ratio. This information will be computerized, allowing the coder either to enter the brand name acronym or to request the list of brand name acronyms for margarine. When information is unknown, selections are based on frequency of use in the general population. Coding rules for individual food items regarding amount of salt used in preparation and at the table will be incorporated into the software development. The interactive dialog for entering this information is shown in Figure 7. The system will be flexible enough so that for studies not interested in sodium intake, the salt portions of the interactive dialog can be bypassed.

Figure 2. Example of documentation in a dietary record.

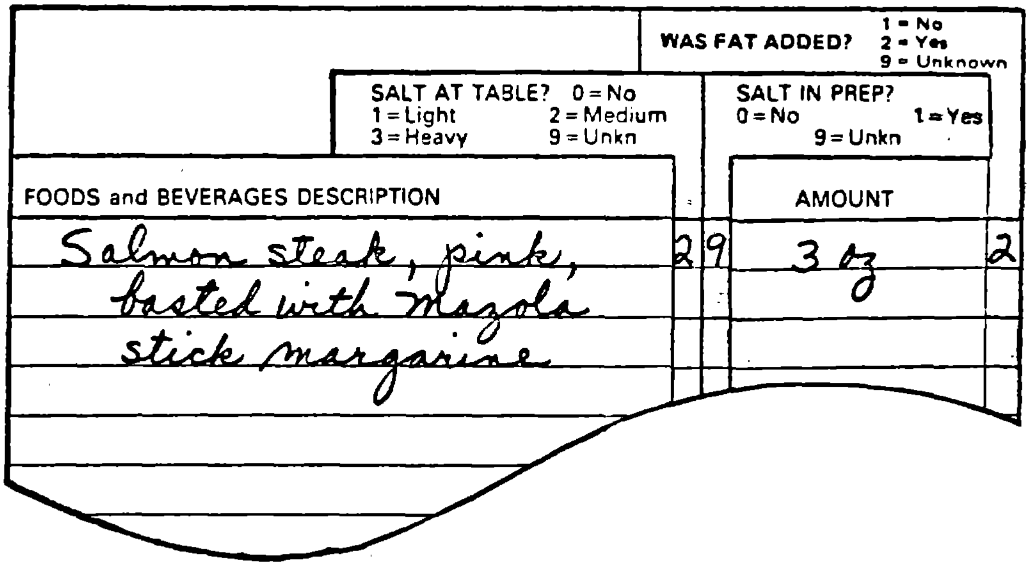
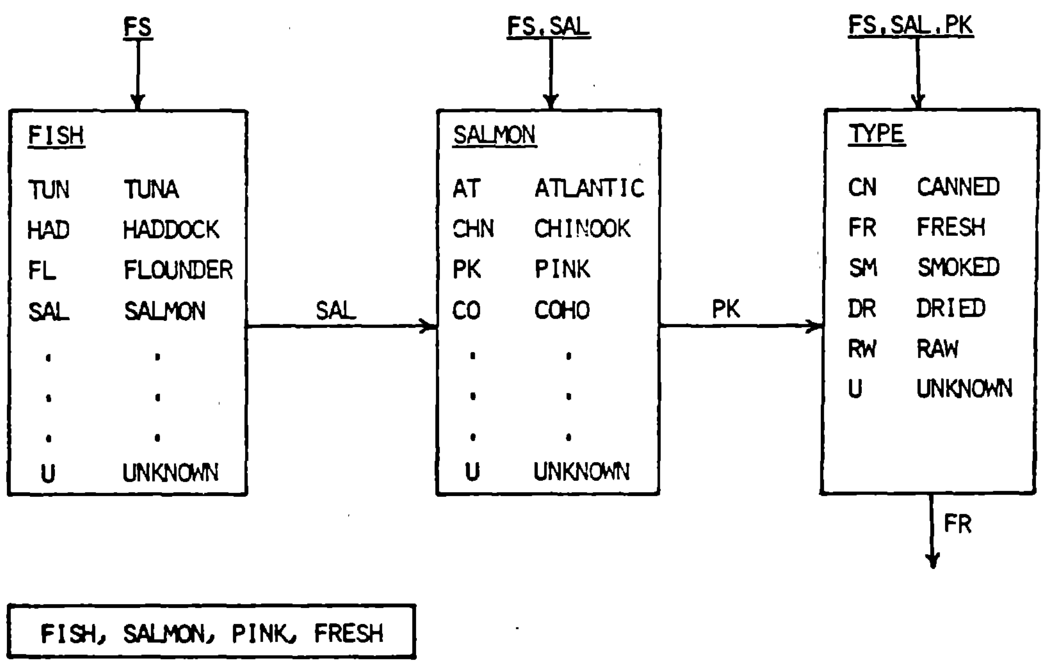


Figure 3. Example of interactive dialog for food item identification.



Note: Information in the boxes appears on the screen in response to information entered by the coder.

Figure 4. Example of interactive dialog for amount consumed.

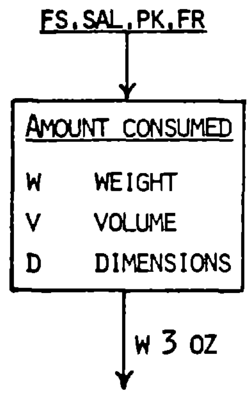


Figure 5. Example of interactive dialog for method of preparation (1).

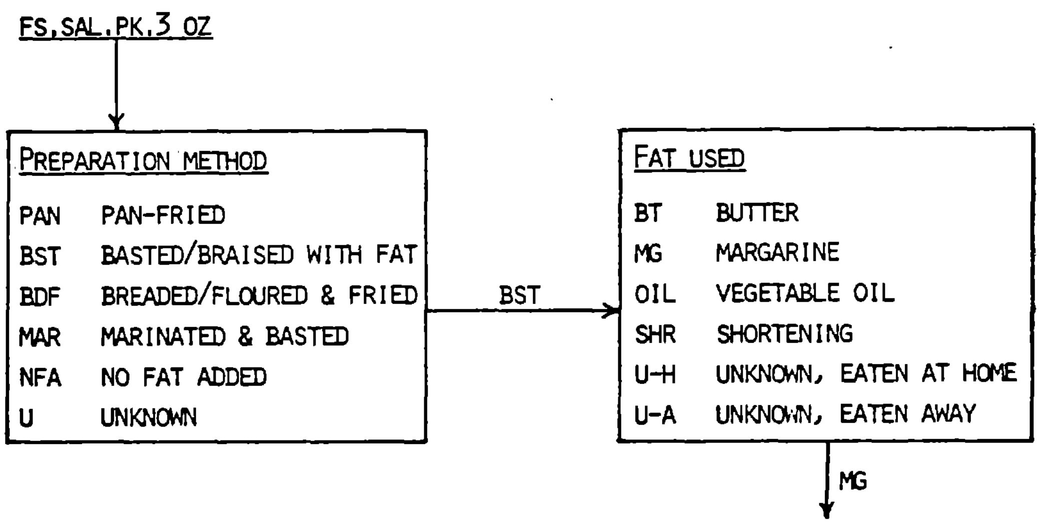


Figure 6. Example of interactive dialog for method of preparation (2).

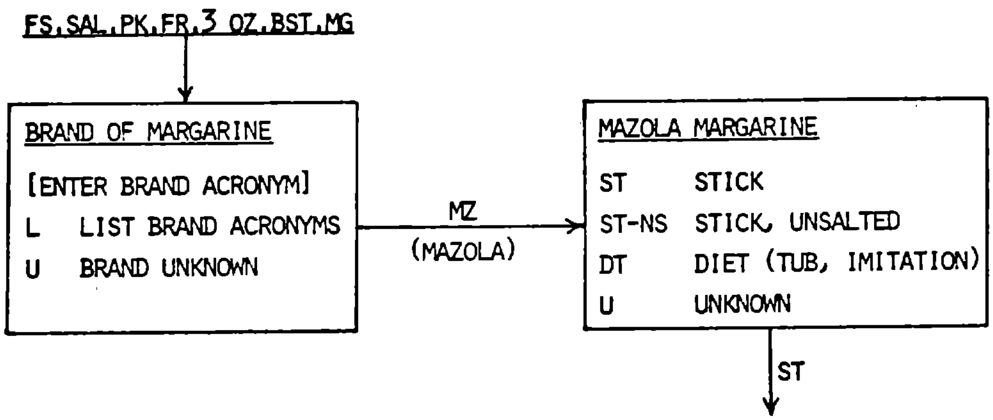
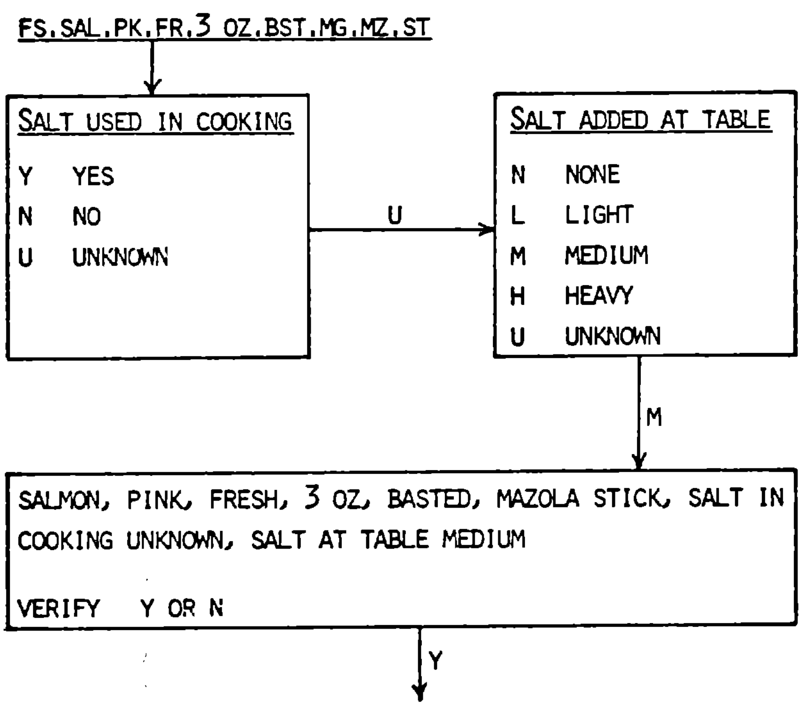


Figure 7. Example of interactive dialog for use of salt.



Once the food item description and preparation have have been specified in detail, the computer will assign the appropriate NCC food codes and quantities. This information will be stored on a disk file for subsequent transmission to a mainframe computer for nutrient calculations and food group analysis.

A feasibility project to demonstrate major aspects of the CNCC model is being developed on the Apple II microcomputer. Although the Apple II does not provide adequate disk storage for the comprehensive system, it is adequate for demonstration purposes. The program for entering the food grouping hierarchies is described in detail by Ellis et all (1983). Software for incorporating preparation codes, salt additions, coding rules, unit conversions, and various edit checks is presently being developed.

EXPANSION-OF THE CNCC SYSTEM TO INCLUDE CODING FOR MULTIPLE NUTRIENT DATA SETS

The CNCC model allows for the addition of food codes for any data set desired. Code numbers for each data set will be added at the end of each interactive pathway. The user will then specify which data set(s) is (are) desired for nutrient calculation, and the computer will assign the appropriate code numbers and amount units. Plans are being developed for incorporating the USDA Nutrient Data Base for Standard Reference (USDA, (1982) into the CNCC system along with the NCC nutrient data set. The USDA data base includes approximately 4,500 items compared with approximately 1,500 items in the NCC data base which was constructed to collapse on similarities in fat content. Based on the maximum amount of descriptive detail documented for any food item, the computer will assign the appropriate codes to either or both of the data sets. A one-to-one correspondence between the two data sets is not necessary as can be seen in the example shown below.

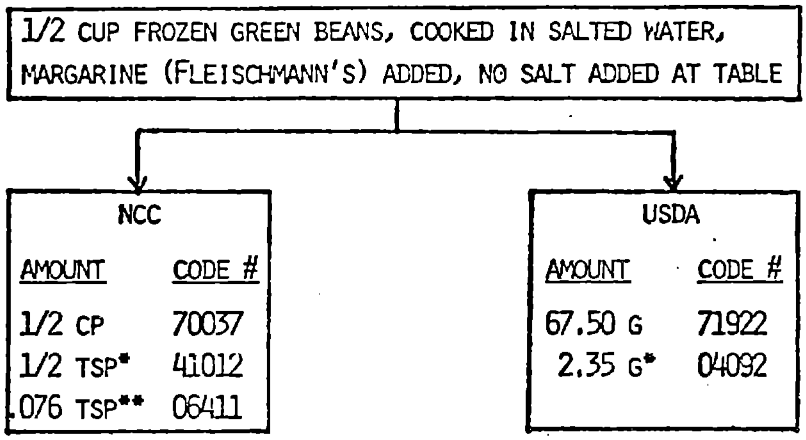
Consider an item documented as follows:

> 1/2 cup frozen green beans, cooked in salted water, margarine (Fleischmann's) added, no salt added at table.

The NCC data base includes only two entries for green beans; one entry is for canned green beans and the other entry includes all other forms. The USDA data set includes 17 green bean entries specifying whether canned green beans are drained or not, diet or regular pack, and whether frozen green beans are cooked or raw, French or regular cut, cooked with or without salt, etc. The documented information will be appropriately used by the computer for assigning both the NCC and the USDA code numbers as indicated in Figure 8. Only two code numbers are required for the USDA data set since salt is included in the USDA green bean code, whereas the salt must be added separately in the NCC system. Coding rules built into the software will specify amounts or types whenever maximum detail is not documented, such as the amount of margarine used in the example above. Since the USDA nutrient calculation system requires input in grams, the computer will generate the appropriate conversion when household or other units are entered.

Figure 8. Example of concurrent coding for more than one nutrient data base.

DOCUMENTATION OF FOOD ITEM:



*Amount based on NCC "prep code" rule for unknown amount of fat added to vegetable

**Amount based on NCC coding rule for salt added in preparation

Code # descriptions:

NCC:	70037	beans, green, ckd, fresh or frozen
	41012	margarine, P/S 2.0-2.5, stick or tub
	06411	salt
USDA	71922	beans, snap, green, frz, cut ckd, boiled, drained, w/salt
	04092	margarine, soft, corn, hydr or reg

Providing the capability of calculating nutrients using various data bases will enable studies to compare their results with studies using other data bases without having to recode the dietary records.

SUMMARY

A model for a microcomputer-based interactive system for coding dietary data is described. The computerized model will incorporate all features of the present NCC coding system including coding rules, methods of food preparation, special procedures for sodium coding, and procedures for verification, error checking, and quality control. The system will be expanded to incorporate coding for multiple nutrient data sets, facilitating comparisons among studies using different data bases.

LITERATURE CITED

Dennis B., N. Ernst, M. Hjortland, J. Tillotson and V. Grambsch, 1980. The NHLBI nutrition data system. J. Am. Dietet. Association. 77:641.

Ellis, L.M.B., S. Chow and M. Buzzard, 1983. Microcomputer-based food coding. Proceedings of the Seventh Annual Symposium on Computer Applications in Medical Care, in press.

Evans, S.N. and A. Gormican, 1973. The computer in retrieving dietary history data. I. Designing and evaluating a computerized diabetic dietary history. J. Am. Dietet. Association. 63:397.

Slack, W., B. Porter, J. Witschi, M. Sullivan, R. Buzbaum and F.J. Stare, 1976. Dietary interviewing by computer. J. Am. Dietet. Association. 69:514.

Tillotson, J.L., D.D. Gorder and N. Kassim, 1981. Nutrition data collection in the Multiple Risk Factor Intervention Trial (MRFIT). J. Am. Dietet. Association. 78:235.

USDA, 1982. Nutrient Data Base for Standard Reference, Release 2.

Witschi, J., H. Kowaloff, S. Bloom, and W. Slack, 1981. Analysis of dietary data: an interactive computer method for storage and retrieval. J. Am. Dietet. Association. 78:609.

THE HIGHS AND LOWS OF BUILDING A DATA BASE

Judith E. Brown, R.D., MPH, Ph.D.
Director and Associate Professor
Program in Public Health Nutrition
School of Public Health
University of Minnesota
Minneapolis, MN 55455
and
President, Nutrition Support Headquarters, Inc.

I am here today representing one of many who have taken the road from nutrient data base ignorance to enlightenment. One who has traveled the road that leads us through the highs and lows of building, maintaining, and implementing a computerized, data base baby. As I have traveled this road and learned where the potholes are, I have taken notes. Today I will give you a book report on the notebook's contents in the hopes of saving those who will take a similar road time and trouble.

PASSAGES IN THE LIFE OF A DATA BASE BUILDER

Building a nutrient data base has its own set of passages. I was introduced to nutrient composition in my freshman year in college. Remember the exercise of recording your diet, looking the foods up, painfully converting household units to grams, and so on? I had no doubt that the results were perfect. I ate 12.3 mg iron each day. Then I became a teacher. I gave that assignment. When reports came in with footnote numbers into the calculations, I began to have my doubts about how accurate these calculations were. So, we went to computerized nutrient analysis, the type that required coding of foods and converting amounts into grams, for better results.

This computer solution failed to take us to nutrient nirvana. Coding and converting measures to grams was a terrific hassle. Once foods were coded, the code was frequently not recognized (because it wasn't included) by the computer program. How many grams are in a half cup of grapes, anyway? Why didn't the program include zinc? And why did we have to make so many substitutions, like entering lasagna because ravioli wasn't in the data base? (Students still thought the results appearing on the "hard copy" were perfect. I was pulling my hair out.)

The Dietary Analysis and Assessment System (nicknamed DAS by users of the system) we were compelled to develop, was built out of total frustration with existing systems. Something had to be done about the coding and converting hassles--these were the two main culprits that took the joy out of nutrient analysis. Why not develop a micro-computer program that would accept foods by their common names and everyday, household measuring units? Why not include those nutrients, for which nutrient composition data exists, that represent important nutrition/health relationships? That defined my dream for a nutrient analysis system.

MAKING THE DREAM COME TRUE

With massive amounts of help from a scientific programmer and hardware specialist, and a substantial financial investment, the dream did come true. During the process, it was up to me to assemble and maintain the data base, and up to the others to get it to work according to the plan.

My life has changed since that day in 1979 when I started to systematically collect food composition data. I went from being the mother of two children to being the mother of two children plus one nutrient data base. All three require extensive nurturing and care, and provide learning experiences. I have learned that there is no data base that is without multiple errors, that contains complete data for all of the nutrients you want, and that includes all of the foods that people eat.

LEARNING TO LIVE WITH ERRORS

> "Honey, if you can't live with mistakes, don't get into the nutrient data base business."
>
> (Dr. E.M. Widdowson, 1983)

That there are multiple errors in data bases comes as no surprise when you consider all of the opportunities for errors. In a data base, say of 1400 foods that includes 35 nutrients per food, there are about 196,000 digits, or 196,000 chances for making errors. Humans enter each of the 196,000 digits, and even with multiple checks, it s difficult to guarantee that each entry is correct.

OTHER SOURCES OF ERRORS

In addition to data base entry errors, bloopers creep into data bases principally in two other ways.

I. HUMAN

- A. Calculating Edible Portion
- B. Calculating/Given Amount
- C. Converting Measurement Units
- D. Recording, Entering, Cross-Checking (Eye-Crossing)
- E. Dietary Data Entry

II. METAPHYSICAL (Sic)

- A. Glitches

When researching a data base, you are sometimes presented two different values for the same food. Pick the wrong one, and you have unwittingly entered an error. Here are three sets of examples of the kinds of choices you face when data differs from reference to reference:

Lowfat Yogurt w/Nonfat Milk Solids, 1c. kcal = 143
Yogurt From Whole Milk, 1c. kcal = 141
Parsley, Raw, Chopped, 1 T = 3.5 g
Parsley, Raw, Chopped, 1 T = 10 g
Fryers, Cooked, Fried, Breast, w/o Ribs, 94 g, KCal = 160
Roasters, Roasted, Light Meat w/o Skin, 94 g, KCal = 171

"WHAT? REALLY? THAT MUST BE WRONG . . ."

Some bloopers can be identified by the outrageous nutrient values assigned. The ones I have kept notes on include:

Rice Krispies, 1c Phosphorous = 0.39
Beef Stroganoff, Frozen, 200 g, kcal = 86
Hollandaise Sauce, 1/4 c, kcal = 180
Round Steak, Bottom, Lean Only, Broiled, 114 g, Riboflavin = 376 mg
Sirloin Steak, Lean, 125 g, Thiamin = 125 mg
Beef Enchiladas, Homemade, 200 g, Vitamin A = 6000 IU

NOTING WHAT'S SAID IN FOOTNOTES

Some of the errors that find their way into data bases can be traced back to their source--the footnotes that are tucked away in small print at the bottom of the pages of nutrient composition tables. Here are three examples of getting tripped by footnotes:

[2]Source of data does not indicate whether raw or cooked, assumed values are for raw food.
[138]Based on total contents of can. If bones are discarded, value will be greatly reduced.
[175]Values range from 60mg to 1,000mg per 100g

USDA MEETS MANUFACTURER'S DATA

Another component of the nutrient data base business is the collection of manufacturer's data. Manufacturer's data are particularly useful because they represent nutrient composition information for foods and brand names that often cannot be found elsewhere. But how does manufacturer's data compare with USDA data when both sets of information are available? My assumption that they would be comparable fell short of the truth:

USDA and Manufacturer's Data
Ex: Cranberry Juice Cocktail, 6 fl. oz.

	Ocean Spray (1978)	USDA #456 (1975)	% Difference USDA/Ocean Spray
Calories	105.7	124	15
Protein	.19 g	.2 g	0
Fat	.19 g	.2 g	0
CHO	26.4 g	31.4 g	16
Calcium	6 mg	10 g	40
Phosphorus	2.3 mg	6 mg	62
Potassium	34.4 mg	19 mg	81
Sodium	3.0 mg	2 mg	50
Iron	.3 mg	.6 mg	50
Thiamin	.02 mg	.02 mg	0
Riboflavin	.03 mg	.02	50
Niacin	.09 mg	.1 mg	0
Vitamin C	60 mg	30 mg	100

Range (0-100%)
= 36% Difference USDA/Ocean Spray

MISSING VALUES ARE A SOURCE OF ERROR

The final source of error that needs to be mentioned is that of missing values. Missing values are to nutrient data bases what the boll weevil is to cotton farmers. Many data base developers boast of the large number of food items and nutrients included in their data bases. What isn't mentioned is the percent of missing values that automatically come with these high figures:

Missing Values*

	Total Foods n	Sample n	No. Food Components	% Missing
USDA #8 (1963)	2483	550	14	13
Bowes & Church's (1980)	4768	450	26	49

*Based on systematic sampling by page number

The question of the extent to which accuracy is lost when a data base is replete with missing values has not been studied. One can only suspect that significant underestimations for particular nutrients results.

Missing data is a particular nemesis when a nutrient/health risk is suspected but data on the particular nutrient is lacking. Data on beta-carotene and simple sugars composition of foods are very incomplete; making it difficult to study the relationships between the nutrients and the incidence of disease. For now, besides undertaking the needed food analyses, we have to live with the data we have. Or, as Margaret Moore so aptly described the situation with simple sugars data:

> "Well, you know, you can't make a silk purse out of a sow's ear".
>
> (personal communication, 1982)

REALITIES OF NUTRIENT DATA BASE IMPLEMENTATION

The final chapter in my notebook was a cause of consternation. Now that my sense of humor has been revived, it is a cause for chuckles. The chapter notes experiences in the real life implementation of DAS. My concerns about the accuracy and completeness of DAS were overshadowed by the realities of dietary data collection. While we wring our hands over the difference in thiamin composition of french cut vs. uncut green beans, there is a health professional out there submitting food records to us on behalf of patients.with entries such as the below:

Food Record Quotes

Food/Description	Amount
Turkey Salad	Large paper plate full
Potato Chips	1 handful
Peanut Butter Bar	1 (stole 2 or 3 off food cart)
Bologna Sandwich	1 bite
Hamburger Casserole (Hamburger, noodles, tomatoes, garlic, cheddar cheese)	1 C.
Brownie	1 - 2" x 1"
Fruit	4 oz.
Meat	2 T.
Cutter's Insect Repellent	2 sprays

Submitters of food records are not without a sense of humor. One particularly poor food record was submitted along with an attached note from the nutritionist that read:

> "I think she'd take in more nutrients if she ate the printout."

Barosso, G. (personal communication, 1983)

NUTRIENT BLOOPER SNOOPERS

Two major problem areas have been highlighted: the errors and missing values in nutrient data bases and the inaccuracies in dietary data collection. The latter is a problem for methodologists and implementors of dietary data collection. The former will become less of a problem as new food analyses are performed. With 200 new food products being offered each month, and other products becoming history, maintaining data bases will forever be with us. The error creep in data bases might be alleviated, in part, if users of nutrient data bases cooperated in the identification of errors. Thus the Nutrient Blooper Snoopers are born. Rewards of, say, one dollar per identified error could be offered. T-Shirts proclaiming "I'm a Nutrient Blooper Snooper" could be presented to those who found the culprits. The point is that, because consumers of nutrient data generally hold the utmost trust in the analysis results, it is incumbent upon us to make data bases as accurate as is humanly possible.

NUTRIENT DATA BASES IN COMPUTER PROGRAMS FOR UNDERGRADUATE FOOD SCIENCE AND NUTRITION CURRICULUM

E. H. Asp
Assistant Professor
Department of Food Science and Nutrition
University of Minnesota

Computer programs for undergraduate education that utilize nutrient data bases have special characteristics. From the program design standpoint, they must have an educational function, besides performing nutrient calculations, the task that computers do quickly, efficiently and accurately. From the program use standpoint, these programs must be easy to use because most student users will be unfamiliar with computers. Programs designed for undergraduate education at the University of Minnesota have not been limited to classroom use; they also have been used in Agricultural Extension work, hospitals, research projects, and the food service industry.

Classes in the Department of Food Science and Nutrition have been using computer programs that incorporate the use of nutrient data bases for the last 15 years. The most interesting changes that have occurred through these years have been the rapid developments in computer technology and the increased interest in determining the nutrient content of food and updating the food composition tables. The computer hardware has changed from the mainframe used in the batch mode in the early years, to another mainframe that was used in the interactive mode for several years, to still another mainframe that is currently used in the interactive mode, to the microcomputers. These changes in the hardware have been accompanied by improvements in the ability of the mainframe computers to handle nutrient data files, especially those in the programs with large data bases. The nutrient data bases in these programs have also had major updates during this time. The most recent was during the last year when the Nutrient Data Base for Standard Reference, Release 2, became the data base for the largest program. The programs developed and used here have changed along with the hardware and nutrient data bases. In fact, the programs are continually being revised and improved.

Computer programs with nutrient data bases are used by undergraduates here in intermediate and advanced Food Science and Nutrition courses. Use of these programs begins in an intermediate-level course, FScN 3272, Introduction to Food Decision Making. In this course, food science subject matter is integrated with nutrition subject matter to cover factors thought by experts to be important considerations when decisions are made about food to be provided for oneself and/or others. Three computer programs are used by students in this course. First, students learn how to use a teletype computer terminal by using a program with a limited food and nutrient data base and only a few commands. Then they use a computer-assisted instructional menu planning program, MENU, which contains a nutrient calculation program that allows the user to revise a problem menu by adding and deleting foods to meet nutrient goals established by the program. Last, students use a more complex program, NUTALLY (Nutrient Tally), with a data base containing over 5000 foods, space for 43 nutrients for each food, and several commands. Students use these programs for projects assigned during the quarter.

Published as Paper No. 1885 of the Miscellaneous Journal Series of the Minnesota Agricultural Experiment Station, St. Paul, Minnesota 55108.

Drawing on the experiences at the University of Minnesota, two aspects of the use of nutrient data bases in computer programs for undergraduate education will be discussed in this paper. They are: (1) program development; and (2) program use by students.

PROGRAM DEVELOPMENT

The question asked most often by those interested in developing a computer program that uses a nutrient data base is "How do I get started?" The next questions usually are "How can I find funding?", and "How are these programs designed and programmed for the computer?"

Getting Started

The first step in program development is to have an idea for a program. Describe the idea in enough detail so an estimate of programming and other costs associated with development and support of the program can be made. This will include investigating computer facilities and capabilities to find the appropriate computer for the program. When developing computer programs that use nutrient data bases, important considerations about the computer are: the storage capacity to handle the data base; the capability for users to manipulate files and update them easily; and quick file access time when retrieving information as the program is being used. These are considerations whether one is working with a mainframe or a microcomputer. Because nutrient data bases tend to be very large, their size may present more storage and access problems when using microcomputers than mainframe computers. The continually increasing storage capacity of microcomputers may eliminate this concern very soon, however. When planning for the use of mainframe computers it is essential to establish contact and cooperation with departmental or university computer personnel at the idea stage to discuss details of programming support available and computer charges for program development and student use.

Funding

Obtaining funds for computer program development at educational institutions requires creativity these days! Funding is often available from a source similar to the Educational Development Program fund at the University of Minnesota that provides grants for projects to improve undergraduate instruction. These grants are awarded to project proposals on a competitive basis. The early program development work here has been funded by Educational Development Program grants. Because some of the programs developed are used in Agricultural Extension work as well as in undergraduate education, and Agricultural Extension is becoming more involved in computer use, some funding has become available from this source. The nutrient data base and program update this past year has received funds from sources outside the university, including a commercial company and a non-university user of one of the programs.

Program Design

When developing computer programs and their nutrient data bases for use in undergraduate courses, objectives for the program and for the course in which the program is to be used must be considered. Programs which are developed to supplement classroom instruction must help students achieve one or more of the objectives for the courses in which they are used. For example, when computer-assisted instructional menu planning program, MENU, was developed, each level of Bloom's Taxonomy of Educational Objectives, Cognitive Domain (Bloom, 1956)--knowledge, comprehension, application, analysis, synthesis and evaluation--was used to develop objectives for the program (Asp and Gordon, 1981).

A content outline for the subject matter to be covered in the program was derived from the course objectives. This outline was used with Bloom's Taxonomy to develop general objectives and specific learning outcomes for each level of the taxonomy using the method designed by Gronlund (Gronlund, 1976).

Selection of the foods and nutrients to include in the nutrient data base depends on the objectives for the program and the characteristics of the students who will use the program. Students starting to learn how to use computer programs find it much easier to work first with programs that use a limited food and nutrient data base; for example, 500 to 600 foods and 11 nutrients for each food, than with programs with large data bases. Students here move to programs with larger data bases (more foods and more nutrients) as their skills in using computer programs improve. Other decisions to be made when developing the food and nutrient data base include: which foods and which forms of the food to use; which unit(s) of measurement to use, e.g., nutrients per serving of food, nutrients per unit of measurement, nutrients per 100 grams edible portion; which nutrients to include; and the source(s) of the nutrient data.

Students entering the "Introduction to Food Decision Making" course here have certain characteristics that must be considered in program design. These students have generally had little or no experience in the use of computers, and many of them have difficulty using math skills. They are unfamiliar with the concept of nutrient data base in a computer program and its relationship to food composition tables. Their knowledge about kinds of food and the preparation and processing terms used to identify food in the food composition tables is limited. These students also are unfamiliar with how food quantities are designated, such as the appropriate serving size for various foods and the conversion of As Purchased quantities to the Edible Portion using percentage refuse data. Their skill in using nutrient names and units of measurements is limited, and they have difficulty identifying food sources of specific nutrients. If students are to become skilled in the use of this information, the computer programs they use must be designed to give them the opportunity to learn and practice these skills.

Computer programs developed for students must be easy for them to use. Students are not always computer experts, yet they are expected to be able to follow program directions and communicate with a program by using its commands. In programs developed here, the directions for using the program, questions, commands, descriptions of the commands and prompts for the user are easy to follow and self-contained in the program, eliminating the need to refer to separate documentation.

Coding of the food items and quantities and the input into the program have been simplified as much as possible, but do vary among the programs the students use. For example, each entry is verified immediately after input in a program students use first to give them practice in checking for accuracy. Students quickly learn that accuracy of the output is no better than accuracy of the input. In programs students use later, an entire list of entries is input, checked for accuracy and the necessary corrections made before continuing with the program.

The output of nutrient data from a computer program must be organized or formatted on the video display or printout so it is easy for the students to read and interpret. Interpretation of the nutrient data from these programs often requires more expertise than is required in the use of computers. In programs students use first, only a few nutrients are included in the output, while in programs used later, the user chooses the number of nutrients desired in the output from several options available. It is important that programs used in education include both (1) nutrient information a professional would want for a client, and (2) options that illustrate various capabilities of computers for displaying the output. The interpretation of numerical output can be enhanced with additional calculations and different types of graphics, depending on the needs of the user. For example, in order to facilitate data interpretations, a calculation and graphing routine is included in programs developed here that prints out and graphs the percentages of a designated Recommended Daily Allowance that is fulfilled by the total for each nutrient for a food list.

The factors just outlined also apply when selecting and using an existing program rather than developing one. Unfortunately, although there appear to be many programs available, an existing program can seldom be found that fulfills all the requirements for use in a new educational setting. It usually must be modified to fit the computer system, the objectives for its use, or in some other way. These modifications can be expensive, therefore, existing programs and their nutrient data bases should be carefully investigated and evaluated before being acquired.

PROGRAM USE BY STUDENTS

Before computer programs with nutrient data bases are used by undergraduate students, it is important to teach the background information that enables students to understand these programs. Because students have had little or no computer experience prior to entering the "Introduction to Food Decision Making" course, they are quite apprehensive about being required to use computer programs for class projects. Their apprehension disappears, however, after they have had the opportunity to learn to use a computer terminal and a program with a limited nutrient data base and only a few commands.

Three questions students frequently ask before using computer programs are: (1) "What is a computer program and what does it do?"; (2) "What is a nutrient data base and how is it used by a computer program?"; (3) "What must I do to be able to use computer programs with nutrient data bases?" The subject matter for the computer unit of the "Introduction to Food Decision Making" course has been organized to answer these questions. The following course activities were developed to provide students with the necessary skills to use computer programs with nutrient data bases as a tool for performing nutrient calculations.

1. Lectures and reading assignments covering basic information about computer programs and how they function, emphasizing use of commands to communicate with programs through computer terminals to input data and receive feedback or output.

2. Lectures describing food composition tables, journal articles and commercial companies as sources of data for nutrient data bases for computer programs.

3. Display of several examples of food composition tables including USDA Agriculture Handbook No. 8 (Watt and Merrill, 1963), its printed updates, sections 8-1 through 8-9, and a printout of the nutrient data from the Nutrient Data Base for Standard Reference, Release 2 computer data tape; USDA Agriculture Handbook No. 456 (Adams, 1975); USDA Home and Garden Bulletin No. 72 (U.S. Department of Agriculture, 1981); and "Bowes and Church's Food Values of Portions Commonly Used" (Pennington and Church, 1980).

4. Reading assignments of two journal articles, one by Hertzler and Hoover (1977) covering use of food composition tables as nutrient data bases for computer programs, and one by Hepburn (1982) describing the USDA national nutrition data bank.

5. Lectures describing the food and nutrient data included in the data bases for the programs students will use for class assignments, followed by a class discussion of the following handouts covering how programs with nutrient data bases operate:

   (a) Directions for using teletype terminals in student computer labs.
   (b) Program printout giving directions for using the MENU program.
   (c) Program printout giving directions and examples of the output for the NUTALLY program.

6. Reading assignments of two journal articles chosen from a list of six articles describing the nutrient calculations performed by different computer programs, each with a different nutrient data base.

7. Problem sets designed to give students practice in finding, calculating and using data to be entered as input into each of the computer programs they will use for class assignments. Problem sets include:

   (a) Designating the recommended serving size in household measures for a list of foods.
   (b) Converting food weights and measures from common units to metric and the reverse.
   (c) Calculating Edible Portion from As Purchased weight and percentage refuse for the foods listed.
   (d) Hand calculating and totaling the amounts of several nutrients in a list of foods as would be done by a computer program.
   (e) Calculating the gram weight for one serving for a list of foods.
   (f) Choosing food codes and specifying food quantities for input into the TALLY and NUTALLY programs.

8. Field trip to the student computer lab where students are supervised by the instructor as they learn how to use computer terminals and programs.

After students learn how to use computer programs with nutrient data bases, they complete two class projects designed to require use of most of the options available in our programs. Students must enter foods into the program (usually foods in a food intake) and then use the program to calculate and print out the nutrient content for each food, the totals for each nutrient for each meal, snack and for the day, percentage of the RDA specified, and graphs of the RDA percentages. The projects also require students to use a computer program to revise a menu to meet specified nutrient goals by deleting and adding foods and adjusting food quantities until the goals are met as indicated by the totals printed out. The student also assumes the role of a professional to interpret the nutrient data in the program output. This interpretation may involve identifying foods that are good or poor sources of specific nutrients, using the data for additional calculations that provide more information, or comparing nutrient totals with recommendations.

SUMMARY

The development and use of computer programs with nutrient data bases for undergraduate education in food science and nutrition provide challenges for food scientists, nutritionists and educators. Describing an idea for a program, obtaining funding and developing an idea into a program is the first challenge. The second is to use the program in an educational setting. To develop basic skills in using computer programs with nutrient data bases, undergraduate students must first learn the background information that enables them to understand these programs and how they are used. After skills in program use have been developed, students should have the opportunity to use and improve these skills by using computer programs for complex projects in advanced courses.

LITERATURE CITED

Adams, C.F. 1975. "Nutritive Value of American Foods in Common Units." Agriculture Handbook No. 456, Agricultural Research Service, United States Department of Agriculture, Washington, D.C.

Asp, E.H. and Gordon, J. 1981. Development of a computer-assisted program for undergraduate instruction. J. Nutr. Educ. 13 Supplement 1:S91-S95.

Bloom, B.C. (ed.) 1956. "Taxonomy of Educational Objectives: Handbook I, Cognitive Domain." David McKay Company, Inc., New York.

Gronlund, N.E. 1976. "Measurement and Evaluation in Teaching." 3rd ed., Macmillan Publishing Company, Inc., New York.

Hepburn, F.N. 1982. The USDA national nutrition data bank. Am. J. Clin. Nutr. 35:1297-1301.

Hertzler, A.A. and Hoover, L.W. 1977. Development of food tables and use with computers. J. Am. Dietet. Assoc. 70:20-29.

Pennington, J.A.T. and Church, H.N. 1980. "Bowes and Church's Food Values of Portions Commonly Used," 13th ed., J.B. Lippincott Company, Philadelphia.

U.S. Department of Agriculture. 1981. "Nutritive Value of Foods." Home and Garden Bulletin No. 72, Consumer Nutrition Center, Science and Education Administration, Washington, D.C.

Watt, B.K. and Merrill, A.L. 1963. "Composition of foods--raw, processed, prepared." Agriculture Handbook No. 8, Consumer and Food Economics Research Division, Agricultural Research Service, United States Department of Agriculture, Washington, D.C.

FUTURE COMPUTERS: NEW PERSPECTIVES
FOR NUTRIENT DATA BANKS

by Earl C. Joseph
President/Futurist
Anticipatory Sciences, Inc.
Minneapolis, Minnesota

ABSTRACT

With the advent of widespread usage of microcomputers and continued rapid advances in computer technology, the future for nutrient data banks is at a turning point pointing towards new directions. This article will characterize this turning point and discuss the implied possible futures opening new perspectives for nutrient data banks.

For example, a breakthrough has occurred in the study of Artificial intelligence which has led to expert systems. As a result, it now becomes possible to design an "Expert Machine"--as an amplifier for giving access to the known and electronically stored knowledge of a field (in the future) to apply in the real-time of one's decisioning, management and control processes. Besides discussing this developing future, a number of other future computer directions and trends will be discussed.

A topical outline of some of the computer futures shaping new perspectives for nutrient banks includes;

- The information age for the 1980's and 90's--a changing environment for society

- Some alternative future information systems--new characteristics, functions and directions

- Artificial intelligence and expert information systems

- Expert systems for amplifying nutrient professionals and users of nutrient data banks

- Imbedded intelligence for automated machines

- Implied nutrient data bank profession turning points, expectations, impacts and consequences for the future in the 1980's, 90's and beyond.

INTRODUCTION

This paper is about the information age and the many trend paths we are on taking us through the decade of the 1980's and toward first quarter 21st Century futures. It discusses both short-term and long-term futures of interest to users of computers--especially for future nutrient data bank users.

In the last decade many breakthroughs were made in our understanding of how the brain works and in the design implementation of artificially-intelligent (AI) systems--especially in the areas of inference leading to expert knowledge-based systems. As a result, we are poised at the brink of a major technology transfer into computer systems of machine intelligence with AI characteristics, algorithms, heuristics and primitives. This paper characterizes possible future computer systems, now in R&D, which implement a number of these AI functions and extrapolates them for the future now being grown for the 21st Century. Included are discussions of computer architectures for the future such as smart machines, future expert systems, knowledge-based architectures and functions, current awareness functions, bio-computers, component systems for decision support, learning, inference discovery, deductive reasoning and problem-solving--and intelligent data bases/banks.

Outlined will be a number of future history maps delineating precursor past trends and alternative possible data processing system futures. Since the sum total of these new directions for the computer field represents a significant breakthrough in what future computer systems will be like and used for in society for assisting humans, this paper also extrapolates these trends and suggests some future impacts to expect.

INFORMATION AGE FUTURES

Massive forces are building for drastically altering which set of alternative futures will become most likely for the remainder of this century--and beyond. Rapid and accelerating advances are occurring in science and technology--especially in computers, communications, artificial intelligence, genetics, agriculture, microbiology, space and chemistry--for altering our future way of life. Economic, political and social forces are also building to cause step function changes.

The emergence of a new information age societal framework, based upon the expanded use of and need for information, supplied by electronic computers and communications, is altering the way in which social, business, economic, educational and political exchanges are conducted. Information age technology is rapidly thundering-in on most jobs and into many homes. Further, since information age technology affects, impacts and alters the way knowledge is created, stored, retrieved and applied, the character and type of jobs we work at, the networks in the way people are linked, the infrastructure of society is changed. These will in turn impact and change the tools we use in the information age, how we use them and for what purposes they will be employed. This is a revolution for society--and in the architecture, organization, processing and dissemination of information/knowledge--for which future computers will perform a central role (but in decentralized embodiments), and will speed the societal transformation underway.

To get a handle on what futures are being grown and what they will be about, one first needs to understand the information age environment being spawned. It has many characteristics and dimensions including an:

Information Ecology, Environment & Sociology
Information Economy and Capitalism
Information Technology, Tools & Systems
Information Resources
Information Conferencing and Dialoguing
Information Politics

Information Ecology includes a new age sociology of many dimensions--a few of which are: information interactions, real-time knowledge access and amplification, information/knowledge engineering, information management, and politics, new job creation and old job displacements, and new individual freedoms and protection options and impacts. In an information age, an information economy transforms capitalism in many forms. Information becomes a basic need and the major social capital as well as societal power source. In an information economy, information further substitutes knowledge for capital, energy, jobs, materials and travel. It de-industrializes jobs by substituting information flows to automations and robotic tools for performing work. An information age thus radically alters societal roles, values, jobs and needs by redesigning the infrastructure of the economy and society.

Information technology marries computers and communications into many other things, and in the process makes "things" smart and intelligent, as well as into our daily lives. Part of this information technology tool kit consists of: people amplifier appliances (computers, calculators, expert systems, etc.), micro-technology/computers/communicators, information appliances (e.g., word processors, "paperless book" systems, computer mail systems, etc.), information systems, information networks, information utilities, information software (DBM/MIS/DSS/OA/CAD/CAM/MRP/KIP/...), data bases and information services. Information technology is part of the information resource environment which additionally includes a rapidly expanding information industry, laws, controls, standards, knowledge bases, media, telecommunication and data communication systems, and computer-aided-systems (e.g., CAD, CAI, CAM, etc.).

In this new information age, people increasingly "discourse" and "dialogue" directly with information via computer terminals, tele- and computer-conferencing and the like, instead of just with people. Information conferencing allows us to conquer distance via electronic computer/communication networks with information appliances backed-up with information services, data bases, knowledge bases and information retrieval and information management software systems. Thus, in such an information environment, policy making, management decision making, people-to-people, people-to-machine, and machine-to-machine interactions, and machine operation increasingly are performed remotely via computer networks using screens, voice channels, keyboards and data bases--instead of face-to-face in conference rooms, back-rooms, offices or in factories.

Information politics included individual privacy and freedom considerations relative to "data-basing people", jobs and employment, societal structuring changes, transborder data flows, information taxing, and the like.

In the process, the information economy now growing increasingly moves a major portion of the GNP from being supplied by the sale of products to sales of information services, i.e., sales of information, information data bases (e.g., nutrient data bases), information systems, information networking, and information assistance.

Taken in total, such a future information environment, now birthing, points to a long list of new computer applications--many of which have their precursors in the primitive (current) data base management system, management information systems, decision support systems and computer-aided-design/instruction systems now in place.

Thus, for the computer word, the movement into this neo-modern information age, which we are leaping into, evolutionary advance is the name of the game rather than revolution--even though for society it is a revolution relative to the magnitude of change occurring and the impacts expected and forecastable.

For example, the current computer revolution differs from the early 1950's, when computers and automation were introduced, in several critical aspect domains. First, the computer application arena is changing from

1) large centralized and costly computer hardware, memory and software to distributed, personal and low-cost systems,
2) use of computers by a few large firms, the military and Federal Government to use by individuals at work and at home, as well as being embedded in a wide variety of machines,
3) displacing "low-level" jobs to displacing "high-level" jobs,
4) creating more jobs than displaced toward the real possibility of displacing jobs faster than society can create new ones,
5) automating industrial age systems to automating post-industrial information age systems and jobs, and
6) computer systems that required considerable education and software development in order to apply them to easy to use systems with a vast array of ready to use package software and data bases.

But the major change occurring is that total computer system and usage costs (hardware, software, terminals, memory, communications, etc.) are at the turning point wherein they are diving ever lower in cost at a faster clip than the inflation rate--and at the same time their functionality and applicability are increasing and widening. Thus, computers, software and data bases have passed the threshold of affordability and are entering an era for being involved in our daily lives, almost world-wide, with billions of people already entering such a computer age.

COMPUTER FUTURES

Both technological knowledge and knowledge about the direction we're traveling into the future, and therefore knowledge about most probable futures, often can have a catalytic impact. That is, a small increase in one's knowledge about either can have an amplifying payoff. Therefore, it's cogent to ask what new computer system developments should we expect for the remainder of this decade and on into the 1990's. Are there new breakthroughs or turning points forecastable? Or, will the decade see only continued but rapid evolutionary developments? What are the dynamics of change to anticipate? In answering these questions, this article maps some of the new territory of future computers for amplifying society in the information age--and for amplifying knowledge workers. Its further purpose is for outlining some likely trends and alternatives for computers: the new powerful information age technological driving force of change, that everyone must become literate of in order to capitalize on its opportunities.

In general, each new future generation of computers will continue to have more memory, be of higher speed, have more functions, be complemented with more application software (programs) and data bases, and be of lower cost, as each succeeding future generation hits the marketplace, than earlier systems as in the past, but at a speeded-up rate.

Computers come in various sizes and with widely different capabilities. The latest and newest type is also the smallest--the microcomputers, or Personal Computers (PCs). Microcomputers today cost as little as 100 dollars and can cost more than 5,000 dollars. Ten years from now, when they are more like components rather than like the boxes today, they could cost one-tenth as much. Microcomputers are screen and keyboard, desktop type devices and were introduced in the late 1970 decade.

The next size computer is called a mini-computer. Today, minis cost up to a hundred times as much as micros and are considerably more capable (at least 10 times) in both their speed of computation and data processing power. Minis also have a much higher memory capacity for storing data and programs than their smaller micro cousins. However, micros and PCs are growing in capability with each new wave of hardware technology advance and are forecasted to take over the mini-computer world as it is known today. Mini-computers are typically file cabinet sized devices and were first introduced in the later half of the 1960 decade.

The next scales of computers can be considerably more capable than mini-computers; they are the mainframe macro-computers (in increasing size):

Small scale computers
Medium scale computers
Large scale computers

These computers were first introduced early in the 1950 decade. The large scale computers are typically room-sized and have an average cost of a million dollars or more.

Forecasters predict that future mini-computers will grow in capability, to become super-minis and thus squeeze out of the running the small and medium scale computers, and that later the micros will take over in this range of computing power also. Further, most future nutrient data banks can be forecasted to migrate toward the microcomputers. They further forecast that large scale computers (today's million dollar class computers) will get bigger and smaller and thus also put the squeeze on the small and medium scale computers, as well as future mini-computers.

The largest scale computers are the supercomputers costing a few million dollars each to more than 20 million dollars--depending on their features, capabilities, functions and capacity. For the future, supercomputers are forecasted to grow in capability by about a factor of ten or more per decade--with a possible step-function increase in capability as we move toward knowledge inference processing systems by the 1990's.

What seems to be happening is that computers are in transition from: 1) a collection of boxes as room-full devices (large scale and super scale computers), to 2) a computer in a single box (e.g., like today's mini-computers in a rack or as the desktop PC microcomputers), to 3) computer on a circuit board and to 4) a computer integrated onto a single chip component. Further, they are being distributed in communication and control networks. A similar trend is occurring with computer <u>systems</u> (collections of computers). The transition that computer systems are undergoing is again along multiple paths, including:

1) centralized and distributed networks of de-centralized collections of boxes;

2) computer systems integrated within a box for centralized applications;

3) distributed/embedded board and chip-level computers integrated as a computer system based in a communications network and later,

4) computer systems (collections of computers) integrated on single wafers and/or on single large chips.

Why are there so many different scales of computers? Simply, to supply varying degrees of capability (like horsepower of motors and cars).

Near-term future trends, in the classical larger computers for large-scale business and scientific applications, and for nutrient data banks, are tending also toward multiple directions; i.e., data base-managed architectures, advanced distributed data processor systems, higher levels of circuit integration, multi-microprocessor architectures, network configurations, considerably smarter and easier-to-use data processing systems, more fault-tolerance architectures (later we may even see self-repairable computers and systems, friendlier computers, knowledge-based systems, integrated smart memory systems, more artificial intelligence, and component computers. Later on, microcomputers will take on these characteristics, and in some cases start the trend with such revolutionary new characteristics--in fact, this process has already started.

Additionally, the need for more data communications (bussing) to link the little ones with other little ones and with the big ones will grow--requiring more high-speed communications (e.g., optical fiber) data links--and an increase in computer networks supplying information like a utility service. "Information utilities" should emerge very soon delivering information (including nutrient data bank information) into homes, offices, schools, labs and factories much like water and electricity are today--thus further widening the usage of computers and PCs.

There are some straightforward reasons why computers and data processing systems are so important today and will continue to be so in the future. First of all, our society is moving into the information age. Computer systems and networks increase the productivity of information activities--it's as simple as that. And, the more capability, the smarter and more intelligent computers become through technology advances, the more that productivity is raised--feeding back to allow growth in our societal quality-of-life.

Software (programs) are on a trend path into the future toward easier-to-use types, more software cast in hardware, growing diversity of applications and lower cost.

Artificial intelligence (AI) type computers, as "inference engines" and "knowledge inference processing" systems are now being researched.

A relatively new technology, Artificial Intelligence (AI) is rapidly advancing beyond the research stage into practical use which soon could be used in nutrient data bank systems. Even though AI has been investigated for about 30 years, it has only been in recent years that AI has been moving toward the development of practical expert systems.

In the last decade we have earned more about the way the human mind works than in all of previous history.

This growth in new knowledge about the human mind and its functioning has lead to recent breakthroughs in the design and implementation of artificially intelligent computer programs. The primary direction of research is especially in the areas of inference reasoning leading to the creation of "expert" knowledge-based systems for amplifying human application of knowledge. As a result, we are poised at the beginning of a major technology transfer into future computer programs and computer systems using artificially intelligent characteristics, algorithms, heuristics, and primitives that, heretofore, were only performable by human minds--via the creation of "expert systems"--for assisting all professions.

But what is an expert system? Today it is a computer system consisting of a set of AI programs that use a stored knowledge base and inference procedures to solve problems. Artificial Intelligence research is a subfield of computer science that investigates the limitation of human processes (within computer systems). These AI processes are called heuristics. Heuristics include earning, intuition, symbolic reasoning, logic rules, inductive discovery and reasoning, deductive analysis, problem solving, and other human intelligence processes including machine representation of knowledge for use in inference tasks. AI assumes such heuristic knowledge is of equal or greater importance than factual knowledge. In fact, for AI purposes, heuristics is assumed to be the process defined as "expertise"--i.e., what "experts" do.

Heuristics goes beyond the use of just logical procedural oriented strings of instructions operating on streams of data, or on data bases--like programs that occur in standard computer systems. Simply, AI heuristics imitates the human brain, especially including heuristics processes for discovering how to solve a problem, or to diagnose.

Conventional information processing computer systems execute a string of instructions (the program), as it streams from memory, processing and transforming data in its memory.

In knowledge processing computer systems, overlayered upon the conventional information processing system, tree strings of knowledge inference procedures working on data, heurstic and logic rules, and question answers are executed from and on information in its knowledge base (memory). Expert systems thread through their knowledge bases via "IF-THEN-AND-ELSE" heuristic rules--that is, the contents of knowledge bases in expert systems is the recodification of knowledge into a form of heuristic logic. Generation of knowledge bases using such logic is akin to programming but without procedure-oriented statements and instructions.

AI heuristics include logical inference procedures which allow semantic access of knowledge bases which use AI processes for making "expert" judgements. AI expert systems capture and store the known expertise of a field (as interpreted from a number of human experts); and translates such knowledge, via AI Programs and hardware that offer intelligent assistance to a practitioner in that field, (i.e., for amplifying a nutrient professional, with its stored knowledge, and AI heuristics for interpreting and applying such expert knowledge). That is, expert systems are people amplifying machines assisting humans in becoming more expert.

Thus, an expert system uses AI inference coupled with a knowledge base for assisting in solving problems, making decisions and judgements, for creating, discovering, planning, designing, or inventing things and opportunities. Expert systems allow the tackling of problems that are difficult enough to require solutions which go beyond simple arithmetic or logic, and that require heuristics of significant power for approaching what heretofore required human experts for their solution. The knowledge and AI heuristic processes necessary to perform at such an expert level, plus the AI inference algorithms used, can be viewed in the AI expert system as a model of the collective expertise of the best human expert practitioners in that field. Knowledge, once captured in such a fashion, in an AI expert system, could also allow a non-expert to apply such expert knowledge and the heuristics to nearly match and often exceed the average unaided human expert in that field. Further, AI expert systems and their knowledge bases can be constantly updated as society gains new knowledge--e.g., education" of expert systems continues via the updating of their knowledge base.

But what is a knowledge base? The process of building a knowledge base for use with an AI expert system requires the compilation of an extremely "factual" taxonomy (including data as now contained in data banks) of the (each) specialized field and the heuristics for its application. Such knowledge-based laxonomies turn out to be far more understandable and accurate, and therefore more useful, than today's manuals and textbooks and data banks. Today expert systems are computer programs employing artificial intelligence operations using knowledge bases for advising people (in an expert fashion) in the "real-time" of the process of doing something.

In forecasting the future of AI expert systems for nutrient data bank usage, there are a number of obvious and expanding application areas. Perhaps at the top of the list for the course of future events for the 1980's is AI advice-giving systems. Already expert system programs exist, or are on the design drawing screens, for medical diagnosis, consulting, architectural design, nutrient expertise, design of very large-scale integrated silicone circuits, molecular generic design, programming, office management decisioning, factory management, home advice (e.g., financial, garden, lawn, repairs), and much more, including applications in the arts. Also being designed are expert management systems, expert programmer systems, and the like. For example, envision how your profession would be enhanced and changed with an expert system or (later) a general purpose mind amplifier.

Future expert systems, in the form of "people amplifiers" (future remote screen and keyboard computers or terminals or hand-held calculator-like devices) present factual data or information and advice or give opinions based upon the AI "knowledge contained in their knowledge bases.

Further, and importantly, an expert system can backtrack to tell the logical process, heuristics, information and logic, what it went through (used) to arrive at its expert advice or opinion.

The knowledge base of an expert system consists of "facts" and heurstics. The "facts" constitute a body or taxonomy of knowledge (and information) that is similar to the information that a human expert would use for whatever expert task such an expert would be performing. But herein lies the stumbling block--what does an expert (human) do?

Further, not all expert knowledge is a set of "black and white" logic facts--much expert knowledge is codifiable only as alternatives, possibles, guesses and opinions (i.e., as fuzzy heuristics). Heuristics, thus, consist of rules of good judgement, fuzzy knowledge, rules of plausible reasoning, as we as hard and fast logical reasoning, rues of good guessing, and the like, that are characteristics of expert-level decision making. Therefore, the performance level of an AI expert system is primarily a function of the speed, size and capacity of its knowledge base, the quality of its contained expert information, the completeness of its taxonomy, and the number and characteristics of its stored or programmed artificial intelligence heuristics (inference rules and procedures).

Today, we are just at the ground floor of creating a variety of knowledge bases. When will/can a variety of knowledge bases be codified for nutrient applications? Only when the nutrient data bank profession gets involved with AI.

The future will see the variety grow and the knowledge base contents evolve with considerable re-codified knowledge. In fact, because a knowledge base arranges knowledge in a somewhat procedural fashion, like a computer program, it must be more complete, correct and comprehensible than the typical textbook or manual. Therefore, experience with current expert system shows that when compared with traditional sources of knowledge (books, tapes, classrooms, data banks, etc.) present and future knowhow based systems are (or can become) 10 to 1,000 times more complete, precise, correct and comprehensible. But perhaps more importantly; as was stated earlier, AI expert systems allow knowledge application in the real-time of human decision making and actions--to amplify humans in real-time decision process.

Is there an AI Expert System in your future? We can now forecast a positive yes.

CONCLUSION

Taken as a whole, these trends allow for both the long and short-term future reality of going beyond science fiction--allowing "Star Trek" like computers and communications systems--and next steps toward (artificial intelligence) "inference engine"-type computers.

Rapid, technological change always has been the norm in the computer field and, recently, in the bio-genetic and communications fields. In the past, technology-driven change has forced an increasing diversity; and from the foregoing, we now see that the same technological-advance trends are forcing the merger of some of this diversity. However; most computer future watchers see this merger as a way for making way for a new form of diversity. The most likely form that such future spintering, now forecastable, will take is among "smart"/"intelligent" vs. "dumb" computer lines, and along application areas.

In summary, this article outlined digital computer and communication trends relative to expectations for: smart computers, convivial systems, current-awareness systems, communicating and discoursing machines, smart computer machines and systems, and artificially intelligent computers of interest to nutrient data bank users and designers.

PC microcomputers are thus evolving into the future along a number of basic paths. One direction of this evolution is its expansion of capabilities--as it chases up the scales of larger computers and displaces some in the process. Another direction, and future for microcomputers, comes about from hardware developments allowing them to be made physically smaller and smaller. This allows them to be embedded into all manner of machines for making the machines they become an integrate part of ever more smarter and (later) more intelligent. When microcomputers are made artificially intelligent then we will call them either "inference engines" or "knowledge inference processing systems".

But, whichever multiple directions the computer and communication fields take into the future, there should be little doubt that these developments will allow computer systems to penetrate deeper into society, providing and making new opportunities and causing considerable change and impact--for providing opportunities for individuals; institutions and society--and most important for helping nutrient data bank users.

FUTURE COMPUTERS: NEW PERSPECTIVES FOR NUTRIENT DATA BANKS

by Earl C. Joseph
President/Futurist
Anticipatory Sciences, Inc.
Minneapolis, Minnesota

ABSTRACT

With the advent of widespread usage of microcomputers and continued rapid advances in computer technology, the future for nutrient data banks is at a turning point pointing towards new directions. This article will characterize this turning point and discuss the implied possible futures opening new perspectives for nutrient data banks.

For example, a breakthrough has occurred in the study of Artificial intelligence which has led to expert systems. As a result, it now becomes possible to design an "Expert Machine"--as an amplifier for giving access to the known and electronically stored knowledge of a field (in the future) to apply in the real-time of one's decisioning, management and control processes. Besides discussing this developing future, a number of other future computer directions and trends will be discussed.

A topical outline of some of the computer futures shaping new perspectives for nutrient banks includes;

- The information age for the 1980's and 90's--a changing environment for society
- Some alternative future information systems--new characteristics, functions and directions
- Artificial intelligence and expert information systems
- Expert systems for amplifying nutrient professionals and users of nutrient data banks
- Imbedded intelligence for automated machines
- Implied nutrient data bank profession turning points, expectations, impacts and consequences for the future in the 1980's, 90's and beyond.

INTRODUCTION

This paper is about the information age and the many trend paths we are on taking us through the decade of the 1980's and toward first quarter 21st Century futures. It discusses both short-term and long-term futures of interest to users of computers--especially for future nutrient data bank users.

In the last decade many breakthroughs were made in our understanding of how the brain works and in the design implementation of artificially-intelligent (AI) systems--especially in the areas of inference leading to expert knowledge-based systems. As a result, we are poised at the brink of a major technology transfer into computer systems of machine intelligence with AI characteristics, algorithms, heuristics and primitives. This paper characterizes possible future computer systems, now in R&D, which implement a number of these AI functions and extrapolates them for the future now being grown for the 21st Century. Included are discussions of computer architectures for the future such as smart machines, future expert systems, knowledge-based architectures and functions, current awareness functions, bio-computers, component systems for decision support, learning, inference discovery, deductive reasoning and problem-solving--and intelligent data bases/banks.

Outlined will be a number of future history maps delineating precursor past trends and alternative possible data processing system futures. Since the sum total of these new directions for the computer field represents a significant breakthrough in what future computer systems will be like and used for in society for assisting humans, this paper also extrapolates these trends and suggests some future impacts to expect.

INFORMATION AGE FUTURES

Massive forces are building for drastically altering which set of alternative futures will become most likely for the remainder of this century--and beyond. Rapid and accelerating advances are occurring in science and technology--especially in computers, communications, artificial intelligence, genetics, agriculture, microbiology, space and chemistry--for altering our future way of life. Economic, political and social forces are also building to cause step function changes.

The emergence of a new information age societal framework, based upon the expanded use of and need for information, supplied by electronic computers and communications, is altering the way in which social, business, economic, educational and political exchanges are conducted. Information age technology is rapidly thundering-in on most jobs and into many homes. Further, since information age technology affects, impacts and alters the way knowledge is created, stored, retrieved and applied, the character and type of jobs we work at, the networks in the way people are linked, the infrastructure of society is changed. These will in turn impact and change the tools we use in the information age, how we use them and for what purposes they will be employed. This is a revolution for society--and in the architecture, organization, processing and dissemination of information/knowledge--for which future computers will perform a central role (but in decentralized embodiments), and will speed the societal transformation underway.

To get a handle on what futures are being grown and what they will be about, one first needs to understand the information age environment being spawned. It has many characteristics and dimensions including an:

- Information Ecology, Environment & Sociology
- Information Economy and Capitalism
- Information Technology, Tools & Systems
- Information Resources
- Information Conferencing and Dialoguing
- Information Politics

Information Ecology includes a new age sociology of many dimensions--a few of which are: information interactions, real-time knowledge access and amplification, information/knowledge engineering, information management, and politics, new job creation and old job displacements, and new individual freedoms and protection options and impacts. In an information age, an information economy transforms capitalism in many forms. Information becomes a basic need and the major social capital as well as societal power source. In an information economy, information further substitutes knowledge for capital, energy, jobs, materials and travel. It de-industrializes jobs by substituting information flows to automations and robotic tools for performing work. An information age thus radically alters societal roles, values, jobs and needs by redesigning the infrastructure of the economy and society.

Information technology marries computers and communications into many other things, and in the process makes "things" smart and intelligent, as well as into our daily lives. Part of this information technology tool kit consists of: people amplifier appliances (computers, calculators, expert systems, etc.), micro-technology/computers/communicators, information appliances (e.g., word processors, "paperless book" systems, computer mail systems, etc.), information systems, information networks, information utilities, information software (DBM/MIS/DSS/OA/CAD/CAM/MRP/KIP/...), data bases and information services. Information technology is part of the information resource environment which additionally includes a rapidly expanding information industry, laws, controls, standards, knowledge bases, media, telecommunication and data communication systems, and computer-aided-systems (e.g., CAD, CAI, CAM, etc.).

In this new information age, people increasingly "discourse" and "dialogue" directly with information via computer terminals, tele- and computer-conferencing and the like, instead of just with people. Information conferencing allows us to conquer distance via electronic computer/communication networks with information appliances backed-up with information services, data bases, knowledge bases and information retrieval and information management software systems. Thus, in such an information environment, policy making, management decision making, people-to-people, people-to-machine, and machine-to-machine interactions, and machine operation increasingly are performed remotely via computer networks using screens, voice channels, keyboards and data bases--instead of face-to-face in conference rooms, back-rooms, offices or in factories.

Information politics included individual privacy and freedom considerations relative to "data-basing people", jobs and employment, societal structuring changes, transborder data flows, information taxing, and the like.

In the process, the information economy now growing increasingly moves a major portion of the GNP from being supplied by the sale of products to sales of information services, i.e., sales of information, information data bases (e.g., nutrient data bases), information systems, information networking, and information assistance.

Taken in total, such a future information environment, now birthing, points to a long list of new computer applications--many of which have their precursors in the primitive (current) data base management system, management information systems, decision support systems and computer-aided-design/instruction systems now in place.

Thus, for the computer word, the movement into this neo-modern information age, which we are leaping into, evolutionary advance is the name of the game rather than revolution--even though for society it is a revolution relative to the magnitude of change occurring and the impacts expected and forecastable.

For example, the current computer revolution differs from the early 1950's, when computers and automation were introduced, in several critical aspect domains. First, the computer application arena is changing from

1) large centralized and costly computer hardware, memory and software to distributed, personal and low-cost systems,
2) use of computers by a few large firms, the military and Federal Government to use by individuals at work and at home, as well as being embedded in a wide variety of machines,
3) displacing "low-level" jobs to displacing "high-level" jobs,
4) creating more jobs than displaced toward the real possibility of displacing jobs faster than society can create new ones,
5) automating industrial age systems to automating post-industrial information age systems and jobs, and
6) computer systems that required considerable education and software development in order to apply them to easy to use systems with a vast array of ready to use package software and data bases.

But the major change occurring is that total computer system and usage costs (hardware, software, terminals, memory, communications, etc.) are at the turning point wherein they are diving ever lower in cost at a faster clip than the inflation rate--and at the same time their functionality and applicability are increasing and widening. Thus, computers, software and data bases have passed the threshold of affordability and are entering an era for being involved in our daily lives, almost world-wide, with billions of people already entering such a computer age.

COMPUTER FUTURES

Both technological knowledge and knowledge about the direction we're traveling into the future, and therefore knowledge about most probable futures, often can have a catalytic impact. That is, a small increase in one's knowledge about either can have an amplifying payoff. Therefore, it's cogent to ask what new computer system developments should we expect for the remainder of this decade and on into the 1990's. Are there new breakthroughs or turning points forecastable? Or, will the decade see only continued but rapid evolutionary developments? What are the dynamics of change to anticipate? In answering these questions, this article maps some of the new territory of future computers for amplifying society in the information age--and for amplifying knowledge workers. Its further purpose is for outlining some likely trends and alternatives for computers: the new powerful information age technological driving force of change, that everyone must become literate of in order to capitalize on its opportunities.

In general, each new future generation of computers will continue to have more memory, be of higher speed, have more functions, be complemented with more application software (programs) and data bases, and be of lower cost, as each succeeding future generation hits the marketplace, than earlier systems as in the past, but at a speeded-up rate.

Computers come in various sizes and with widely different capabilities. The latest and newest type is also the smallest--the microcomputers, or Personal Computers (PCs). Microcomputers today cost as little as 100 dollars and can cost more than 5,000 dollars. Ten years from now, when they are more like components rather than like the boxes today, they could cost one-tenth as much. Microcomputers are screen and keyboard, desktop type devices and were introduced in the late 1970 decade.

The next size computer is called a mini-computer. Today, minis cost up to a hundred times as much as micros and are considerably more capable (at least 10 times) in both their speed of computation and data processing power. Minis also have a much higher memory capacity for storing data and programs than their smaller micro cousins. However, micros and PCs are growing in capability with each new wave of hardware technology advance and are forecasted to take over the mini-computer world as it is known today. Mini-computers are typically file cabinet sized devices and were first introduced in the later half of the 1960 decade.

The next scales of computers can be considerably more capable than mini-computers; they are the mainframe macro-computers (in increasing size):

Small scale computers
Medium scale computers
Large scale computers

These computers were first introduced early in the 1950 decade. The large scale computers are typically room-sized and have an average cost of a million dollars or more.

Forecasters predict that future mini-computers will grow in capability, to become super-minis and thus squeeze out of the running the small and medium scale computers, and that later the micros will take over in this range of computing power also. Further, most future nutrient data banks can be forecasted to migrate toward the microcomputers. They further forecast that large scale computers (today's million dollar class computers) will get bigger and smaller and thus also put the squeeze on the small and medium scale computers, as well as future mini-computers.

The largest scale computers are the supercomputers costing a few million dollars each to more than 20 million dollars--depending on their features, capabilities, functions and capacity. For the future, supercomputers are forecasted to grow in capability by about a factor of ten or more per decade--with a possible step-function increase in capability as we move toward knowledge inference processing systems by the 1990's.

What seems to be happening is that computers are in transition from: 1) a collection of boxes as room-full devices (large scale and super scale computers), to 2) a computer in a single box (e.g., like today's mini-computers in a rack or as the desktop PC microcomputers), to 3) computer on a circuit board and to 4) a computer integrated onto a single chip component. Further, they are being distributed in communication and control networks. A similar trend is occurring with computer systems (collections of computers). The transition that computer systems are undergoing is again along multiple paths, including:

1) centralized and distributed networks of de-centralized collections of boxes;

2) computer systems integrated within a box for centralized applications;

3) distributed/embedded board and chip-level computers integrated as a computer system based in a communications network and later,

4) computer systems (collections of computers) integrated on single wafers and/or on single large chips.

Why are there so many different scales of computers? Simply, to supply varying degrees of capability (like horsepower of motors and cars).

Near-term future trends, in the classical larger computers for large-scale business and scientific applications, and for nutrient data banks, are tending also toward multiple directions; i.e., data base-managed architectures, advanced distributed data processor systems, higher levels of circuit integration, multi-microprocessor architectures, network configurations, considerably smarter and easier-to-use data processing systems, more fault-tolerance architectures (later we may even see self-repairable computers and systems, friendlier computers, knowledge-based systems, integrated smart memory systems, more artificial intelligence, and component computers. Later on, microcomputers will take on these characteristics, and in some cases start the trend with such revolutionary new characteristics--in fact, this process has already started.

Additionally, the need for more data communications (bussing) to link the little ones with other little ones and with the big ones will grow--requiring more high-speed communications (e.g., optical fiber) data links--and an increase in computer networks supplying information like a utility service. "Information utilities" should emerge very soon delivering information (including nutrient data bank information) into homes, offices, schools, labs and factories much like water and electricity are today--thus further widening the usage of computers and PCs.

There are some straightforward reasons why computers and data processing systems are so important today and will continue to be so in the future. First of all, our society is moving into the information age. Computer systems and networks increase the productivity of information activities--it's as simple as that. And, the more capability, the smarter and more intelligent computers become through technology advances, the more that productivity is raised--feeding back to allow growth in our societal quality-of-life.

Software (programs) are on a trend path into the future toward easier-to-use types, more software cast in hardware, growing diversity of applications and lower cost.

Artificial intelligence (AI) type computers, as "inference engines" and "knowledge inference processing" systems are now being researched.

A relatively new technology, Artificial Intelligence (AI) is rapidly advancing beyond the research stage into practical use which soon could be used in nutrient data bank systems. Even though AI has been investigated for about 30 years, it has only been in recent years that AI has been moving toward the development of practical expert systems.

In the last decade we have earned more about the way the human mind works than in all of previous history.

This growth in new knowledge about the human mind and its functioning has lead to recent breakthroughs in the design and implementation of artificially intelligent computer programs. The primary direction of research is especially in the areas of inference reasoning leading to the creation of "expert" knowledge-based systems for amplifying human application of knowledge. As a result, we are poised at the beginning of a major technology transfer into future computer programs and computer systems using artificially intelligent characteristics, algorithms, heuristics, and primitives that, heretofore, were only performable by human minds--via the creation of "expert systems"--for assisting all professions.

But what is an expert system? Today it is a computer system consisting of a set of AI programs that use a stored knowledge base and inference procedures to solve problems. Artificial Intelligence research is a subfield of computer science that investigates the limitation of human processes (within computer systems). These AI processes are called heuristics. Heuristics include earning, intuition, symbolic reasoning, logic rules, inductive discovery and reasoning, deductive analysis, problem solving, and other human intelligence processes including machine representation of knowledge for use in inference tasks. AI assumes such heuristic knowledge is of equal or greater importance than factual knowledge. In fact, for AI purposes, heuristics is assumed to be the process defined as "expertise"--i.e., what "experts" do.

Heuristics goes beyond the use of just logical procedural oriented strings of instructions operating on streams of data, or on data bases--like programs that occur in standard computer systems. Simply, AI heuristics imitates the human brain, especially including heuristics processes for discovering how to solve a problem, or to diagnose.

Conventional information processing computer systems execute a string of instructions (the program), as it streams from memory, processing and transforming data in its memory.

In knowledge processing computer systems, overlayered upon the conventional information processing system, tree strings of knowledge inference procedures working on data, heurstic and logic rules, and question answers are executed from and on information in its knowledge base (memory). Expert systems thread through their knowledge bases via "IF-THEN-AND-ELSE" heuristic rules--that is, the contents of knowledge bases in expert systems is the recodification of knowledge into a form of heuristic logic. Generation of knowledge bases using such logic is akin to programming but without procedure-oriented statements and instructions.

AI heuristics include logical inference procedures which allow semantic access of knowledge bases which use AI processes for making "expert" judgements. AI expert systems capture and store the known expertise of a field (as interpreted from a number of human experts); and translates such knowledge, via AI Programs and hardware that offer intelligent assistance to a practitioner in that field, (i.e., for amplifying a nutrient professional, with its stored knowledge, and AI heuristics for interpreting and applying such expert knowledge). That is, expert systems are people amplifying machines assisting humans in becoming more expert.

Thus, an expert system uses AI inference coupled with a knowledge base for assisting in solving problems, making decisions and judgements, for creating, discovering, planning, designing, or inventing things and opportunities. Expert systems allow the tackling of problems that are difficult enough to require solutions which go beyond simple arithmetic or logic, and that require heuristics of significant power for approaching what heretofore required human experts for their solution. The knowledge and AI heuristic processes necessary to perform at such an expert level, plus the AI inference algorithms used, can be viewed in the AI expert system as a model of the collective expertise of the best human expert practitioners in that field. Knowledge, once captured in such a fashion, in an AI expert system, could also allow a non-expert to apply such expert knowledge and the heuristics to nearly match and often exceed the average unaided human expert in that field. Further, AI expert systems and their knowledge bases can be constantly updated as society gains new knowledge--e.g., education" of expert systems continues via the updating of their knowledge base.

But what is a knowledge base? The process of building a knowledge base for use with an AI expert system requires the compilation of an extremely "factual" taxonomy (including data as now contained in data banks) of the (each) specialized field and the heuristics for its application. Such knowledge-based laxonomies turn out to be far more understandable and accurate, and therefore more useful, than today's manuals and textbooks and data banks. Today expert systems are computer programs employing artificial intelligence operations using knowledge bases for advising people (in an expert fashion) in the "real-time" of the process of doing something.

In forecasting the future of AI expert systems for nutrient data bank usage, there are a number of obvious and expanding application areas. Perhaps at the top of the list for the course of future events for the 1980's is AI advice-giving systems. Already expert system programs exist, or are on the design drawing screens, for medical diagnosis, consulting, architectural design, nutrient expertise, design of very large-scale integrated silicone circuits, molecular generic design, programming, office management decisioning, factory management, home advice (e.g., financial, garden, lawn, repairs), and much more, including applications in the arts. Also being designed are expert management systems, expert programmer systems, and the like. For example, envision how your profession would be enhanced and changed with an expert system or (later) a general purpose mind amplifier.

Future expert systems, in the form of "people amplifiers" (future remote screen and keyboard computers or terminals or hand-held calculator-like devices) present factual data or information and advice or give opinions based upon the AI "knowledge contained in their knowledge bases.

Further, and importantly, an expert system can backtrack to tell the logical process, heuristics, information and logic, what it went through (used) to arrive at its expert advice or opinion.

The knowledge base of an expert system consists of "facts" and heurstics. The "facts" constitute a body or taxonomy of knowledge (and information) that is similar to the information that a human expert would use for whatever expert task such an expert would be performing. But herein lies the stumbling block--what does an expert (human) do?

Further, not all expert knowledge is a set of "black and white" logic facts--much expert knowledge is codifiable only as alternatives, possibles, guesses and opinions (i.e., as fuzzy heuristics). Heuristics, thus, consist of rules of good judgement, fuzzy knowledge, rules of plausible reasoning, as we as hard and fast logical reasoning, rues of good guessing, and the like, that are characteristics of expert-level decision making. Therefore, the performance level of an AI expert system is primarily a function of the speed, size and capacity of its knowledge base, the quality of its contained expert information, the completeness of its taxonomy, and the number and characteristics of its stored or programmed artificial intelligence heuristics (inference rules and procedures).

Today, we are just at the ground floor of creating a variety of knowledge bases. When will/can a variety of knowledge bases be codified for nutrient applications? Only when the nutrient data bank profession gets involved with AI.

The future will see the variety grow and the knowledge base contents evolve with considerable re-codified knowledge. In fact, because a knowledge base arranges knowledge in a somewhat procedural fashion, like a computer program, it must be more complete, correct and comprehensible than the typical textbook or manual. Therefore, experience with current expert system shows that when compared with traditional sources of knowledge (books, tapes, classrooms, data banks, etc.) present and future knowhow based systems are (or can become) 10 to 1,000 times more complete, precise, correct and comprehensible. But perhaps more importantly; as was stated earlier, AI expert systems allow knowledge application in the real-time of human decision making and actions--to amplify humans in real-time decision process.

Is there an AI Expert System in your future? We can now forecast a positive yes.

CONCLUSION

Taken as a whole, these trends allow for both the long and short-term future reality of going beyond science fiction--allowing "Star Trek" like computers and communications systems--and next steps toward (artificial intelligence) "inference engine"-type computers.

Rapid, technological change always has been the norm in the computer field and, recently, in the bio-genetic and communications fields. In the past, technology-driven change has forced an increasing diversity; and from the foregoing, we now see that the same technological-advance trends are forcing the merger of some of this diversity. However; most computer future watchers see this merger as a way for making way for a new form of diversity. The most likely form that such future spintering, now forecastable, will take is among "smart"/"intelligent" vs. "dumb" computer lines, and along application areas.

In summary, this article outlined digital computer and communication trends relative to expectations for: smart computers, convivial systems, current-awareness systems, communicating and discoursing machines, smart computer machines and systems, and artificially intelligent computers of interest to nutrient data bank users and designers.

PC microcomputers are thus evolving into the future along a number of basic paths. One direction of this evolution is its expansion of capabilities--as it chases up the scales of larger computers and displaces some in the process. Another direction, and future for microcomputers, comes about from hardware developments allowing them to be made physically smaller and smaller. This allows them to be embedded into all manner of machines for making the machines they become an integrate part of ever more smarter and (later) more intelligent. When microcomputers are made artificially intelligent then we will call them either "inference engines" or "knowledge inference processing systems".

But, whichever multiple directions the computer and communication fields take into the future, there should be little doubt that these developments will allow computer systems to penetrate deeper into society, providing and making new opportunities and causing considerable change and impact--for providing opportunities for individuals; institutions and society--and most important for helping nutrient data bank users.

COMPUTER SYSTEMS

James J. Doyle, Honeywell, Datanetwork
Loretta W. Hoover, University of Missouri-Columbia

QUESTIONS--COMPUTER SYSTEMS

<u>SECTION 1</u>:

1. File--Average size of data record for a nutrient; size of name minimum for data compression; data storage techniques.

2. Future applications on pocket computer.

3. Future computers from users.

4. Languages being used for access to data bases.

5. Interface micros to man phones x.25.

6. Developing computer literacy among the uninformed.

<u>SECTION 2</u>:

1. Noncompatibility between computer systems.

2. Clearing house for computerized food service problems.

3. Other meetings of conferences deserve with hardware/software.

4. Imbed nutrient data into a comprehensive system.

5. Limitations of micros? How long should I wait?

6. How do you continue your current level while converting to a computer system?

7. What interest from the general public about nutrition information?

8. How does a novice select equipment and software?

9. Legal liabilities of computers.

<u>SECTION 3</u>:

1. Calculated vs. analytical data--any correspondence.

2. Specific variations for Vit A and Vit C.

3. Varying disciplines.

DATA BASE DEVELOPMENT

Lillian White, Dietary Analysis and Prescription Services, Inc.
Jean Pennington, Food and Drug Administration

QUESTIONS DISCUSSED:

1. Recipe calculation programs:

   Use of USDA Hnd 102 to determine moisture/fat loss or gain (dry macaroni, etc.).

   Data from Hnd 102 is computerized and adjusted into food codes in the data base.

2. Use of program to evaluate nutrient density and % RDA.

3. Composition of water; use of water softeners; effect of mineral content of water on mineral intake.

4. Availability of data on allergens, caffeine and other non-nutrients.

5. Documentation for nutrient values; public domain of values; copyright of data.

6. Clearing house for nutrient data.

7. Problem of freezing data at one point in time.

8. How to input values.

9. How to incorporate data from different sources.

EDUCATION

Elaine H. Asp, University of Minnesota
Joanne L. Slavin, University of Minnesota

To begin the session, we discussed how computers are being used in various educational settings in Minnesota. The use of the computer in informal nutrition education was presented, as was how the computer has been incorporated into the public school curriculum, for elementary and junior and senior high students. Uses of the computer in the training of foods and nutrition students was presented. Other health professions, such as nursing, medicine, and dentistry, are also using foods and nutrition programs in college courses and these were presented by different attendees of the session.

Various microcomputers and software were used in shopping mall settings as informal nutrition education for the public. CEREALS, a simple program distributed by the Minnesota Educational Computing Consortium (MECC), was recommended for use with the public. The program shows the user graphically how much sugar is in different cereal products. The user does not learn anything about the nutritional attributes of sugar. In other words, users will leave the program knowing that their favorite cereal is 40% sugar, but they will not know what to do with that information.

The use of Pillsbury's "Eat Smart" Apple II program at the Twin Cities Marathon was discussed, Runners were very interested in having their diets analyzed, but the nutrient analysis generated many technical nutrition questions. Fortunately, nutrition graduate students were available to answer questions. However, the experience illustrated that nutrition computer programs will not replace nutritionists, but may change their responsibilities.

Elementary school students are some of the biggest users of foods and nutrition computer programs. Some of the workshop participants described their efforts in teaching students how to use foods and nutrition programs. MECC software has been used, as has software from the University of California-Berkeley called "What's in your lunch". Games are very popular vehicles for teaching foods and nutrition concepts.

Computer camps for elementary children that included foods and nutrition programs have been conducted by some workshop participants. The Dairy Council is developing software on foods and nutrition for elementary children which should be available soon. There was some discussion about access to home computers. Obviously, some children only have access to a computer at school or camp. Educators should be sensitive to this discrepancy and not expect all children to be equally adept at using the computer.

The uses of computer in professional education was discussed during both sessions. Although there was general agreement that foods and nutrition students should have experience with computer, some programs make minimal use of the computer in professional courses. Also, it was discussed whether a separate course on computers should be developed or if computer technology should be incorporated into existing courses. Schools, such as Texas Women's University, do have a separate computer course for dietetic students. The course is team taught with faculty from both computer science and nutrition.

Although most of the groups agreed such a course would be advantageous, most universities are not at liberty to develop new courses, because of fiscal restraints. Therefore, most faculty members present described how computers have been or will be incorporated into existing foods and nutrition courses. In some schools, such as Eastern Missouri University, all graduates must be "computer literate", putting pressure on dietetics faculty to include computer concepts in their teaching. About half of the universities represented had microcomputer labs for their students, while the rest had students working on the mainframe systems. Most participants were anxious to get more microcomputers for their students to use before incorporating microcomputers into their classes.

Some time was spent discussing drawbacks to computers. Participants pointed out that computer printouts tend to be gospel, even if errors in the data base or calculation logic give gross errors. When diet analysis programs are used, students may not learn how to calculate the percentage of the RDA, if the computer does the work for them. It was agreed that students should be required to interpret the results of computer programs used in various classes, not just run the program. Many of the diet analysis programs used in college courses, therefore, may not be "user-friendly" enough for hospital staff or the public.

Directories for software were suggested by various participants. Penn State was identified as a source for a nutrition software directory. Personal Computing, the August 1983 issue lists nutrition and health software for microcomputers. Problems with software quality were debated. The June 1984 issue of the Journal of Nutrition Education will be devoted totally to computers in nutrition education.

Access to computers was a problem for some of the participants. Microcomputer manufacturers were recommended as potential donators of hardware. Aileen Eick, Mankato State University, suggested National Science Foundation (NSF) as a potential funding agency, since her school received a large grant from NSF to incorporate microcomputers into the curriculum.

The session was useful in sharing sources for software and describing successful settings for foods and nutrition computer programs. Some participants had not yet incorporated computers into their teaching situation and found it useful listening to those experienced with computers in nutrition teaching.

INDUSTRIAL USES

Rose Tobelmann, General Mills, Inc.
Barbara Bentson, The Pillsbury Company

Both Barb Benson and Rose Tobelmann presented a short synopsis of the uses of nutrient data bases within their respective corporations, Pillsbury and General Mills.

The discussion following these presentations centered around the nutrient information corporations release to the public for inclusion in various nutrient data banks. Academic representative indicated the basic nutrition label information was not enough. Rather, they need complete nutrient analytical information for the basic vitamins and minerals as well as trace nutrient, total sugar and complete fat analysis.

Industry representatives indicated the cost of these more involved analyses would be phenomenal when one starts to consider the amount of reformulation many products encounter. Additionally, some smaller industries do not have the dollars nor the priority to complete nutrient analysis on their products.

There was also discussion about industry's responsibility to provide nutrient information for research purposes. Although many in the audience felt it was industry's responsibility, the industry's side of the story centers around the uses of the information.

Why should industry release specific nutrient information (that is not required by law to be listed on the nutrition label) i.e., sugar for research purposes when the results of this research could negatively and potentially invalidly implicate the food product.

Other industry representatives in the audience also contributed their corporations policy on releasing nutrient data. Beth Rusnak resounded Frito-Lay's corporate commitment to nutrition and indicated Frito will release 100 gram analytical information to the public. Nancy Rawson reiterated the Campbell Soup Company's position regarding nutrient analysis. Each of the company's products are analyzed and these results are available upon request. Nancy did indicate that some of Campbell's subsidiary's products, i.e., Pepperidge Farm, are not yet analyzed, however, they are working in that area.

A final point of discussion dealt with the large number of product reformulations involved in the food industry. All industry representatives reaffirmed that these reformulations will continue and may potentially increase--that is a consistent factor in the food industry. However, a suggestion was made concerning the development of a clearing house for food industry nutrient information.

With such a system set up, industry would only need to update a source of information and anyone needing nutrient information for a nutrient data base could contact the clearing house and obtain the most current product nutrient information.

PATIENT CARE - DIETARY EVALUATION

Ann T. Sandvick, Children's Hospital, Medical Center, Akron, Ohio
Joyce B. Wenz, University of Minnesota

Both Joyce Wenz and Ann Sandvick gave a brief description of how they have used nutrient data banks in hospital/patient care situation. Listed below are some of the questions the audience raised and for some, the responses recorded. There was significant dicussion concerning the costs involved.

What do you charge for a 24-hour analysis?
$5.00 for a 3-day diary.

More detail on charges please--
Nutritional status assessment
Must have an order for it or must be a standard protocol
Instructions and follow-ups are charged
e.g., diabetic - $30, then $15

Go back to statement about charging for special diets--any objection by patient?
If NPO, not charged for that meal.

No trouble for charging? Don't have to go through pharmacy?
No.

Is there 3rd party reimbursement?
Approximately 80% from major companies for inpatients. For outpatients, it varies.

Is it covered by third party?
Charge for nutrition available
Charge for labor involved in a special diet--not the food
Charge for instruction and special assessments

What companies pay?
Major companies

Staff-patient ratio?
At the Children's Hospital, 4 for 175 patients.

Who puts the data into the system?
Computer technician.

Did you have to change the staff?
Yes--had to upgrade positions.

How many techs?
Two--they do data input.

Who interviews?
Dietitians.

What special training have the dietitians received?
How literate are they in terms of computer work.
    On the job training.
    Jobs have been upgraded.

Our dietitians are resistant to codes. How to handle? How do you handle checks backs?
    Have a technician do that.
    Terminals are available.

Who does the coding?
    The tech.

How long before you see the point where the patient is using terminals?
    Hard to predict, could be a problem on the floors.

Let me understand, the food service worker does the weighing after?
    Yes.

You said you enter from supplements and parential.
    We use a standard nutritional assessment form.

How handle the supplements?
    Subtotal, then add in.

Is the assessment hand written?
    No--dictated.

Do you show the patient the printout?
    Yes, sometimes. We discuss it.

Concerning unknowns
    Know how many we have values for.

Have you used it for a quality control and cost effectiveness?
    Not yet.

Suggest--talk to medical records people for what you want as a diet in terms of DRG's.

How do you update your database and how often?
    Purchase updated annually.

Do you have the ability to make changes (i.e., when you find errors)?
    Yes, that day.

How many calorie intake analyses do you do a day?
    Probably 20.

How does the check-back work?
    The tech does it.

With the experience you have today, what would you do different?
There are more software packages available.
Look carefully and imagine how the system can grow.

Does anyone have a system they like?
Hard disc faster than floppy.

Control--are items weighted and measured?
Yes, if patient is on check-back.

Medical Software Systems, Des Moines uses an optically read ticket--has a database.

Do you do inventory?
No, it is done manually.

What institutions are using it?
Some hospitals are awaiting installation.
Any assessment done by computer?
No.

How do you report "holes" in diet, as a missing food?
There's a space at end of each day for additions and comments.

When you first purchased your database from Case Western, did they give the programs or did you do it yourself?
We did all the programs.

Do you not feel this is more costly as opposed to purchasing a program?
There was not much available.

What kind of summaries are you printing from your program?
Two page summary but a lot <u>incomplete</u>.

Do you have nutritional assessment on the computer?
No.

PATIENT CARE - DIETARY EVALUATION

Ann T. Sandvick, Children's Hospital, Medical Center, Akron, Ohio
Joyce B. Wenz, University of Minnesota

Both Joyce Wenz and Ann Sandvick gave a brief description of how they have used nutrient data banks in hospital/patient care situation. Listed below are some of the questions the audience raised and for some, the responses recorded. There was significant dicussion concerning the costs involved.

What do you charge for a 24-hour analysis?
$5.00 for a 3-day diary.

More detail on charges please--
Nutritional status assessment
Must have an order for it or must be a standard protocol
Instructions and follow-ups are charged
e.g., diabetic - $30, then $15

Go back to statement about charging for special diets--any objection by patient?
If NPO, not charged for that meal.

No trouble for charging? Don't have to go through pharmacy?
No.

Is there 3rd party reimbursement?
Approximately 80% from major companies for inpatients. For outpatients, it varies.

Is it covered by third party?
Charge for nutrition available
Charge for labor involved in a special diet--not the food
Charge for instruction and special assessments

What companies pay?
Major companies

Staff-patient ratio?
At the Children's Hospital, 4 for 175 patients.

Who puts the data into the system?
Computer technician.

Did you have to change the staff?
Yes--had to upgrade positions.

How many techs?
Two--they do data input.

Who interviews?
Dietitians.

What special training have the dietitians received?
How literate are they in terms of computer work.

On the job training.
Jobs have been upgraded.

Our dietitians are resistant to codes. How to handle? How do you handle checks backs?

Have a technician do that.
Terminals are available.

Who does the coding?

The tech.

How long before you see the point where the patient is using terminals?

Hard to predict, could be a problem on the floors.

Let me understand, the food service worker does the weighing after?

Yes.

You said you enter from supplements and parential.

We use a standard nutritional assessment form.

How handle the supplements?

Subtotal, then add in.

Is the assessment hand written?

No--dictated.

Do you show the patient the printout?

Yes, sometimes. We discuss it.

Concerning unknowns

Know how many we have values for.

Have you used it for a quality control and cost effectiveness?

Not yet.

Suggest--talk to medical records people for what you want as a diet in terms of DRG's.

How do you update your database and how often?

Purchase updated annually.

Do you have the ability to make changes (i.e., when you find errors)?

Yes, that day.

How many calorie intake analyses do you do a day?

Probably 20.

How does the check-back work?

The tech does it.

With the experience you have today, what would you do different?
    There are more software packages available.
    Look carefully and imagine how the system can grow.

Does anyone have a system they like?
    Hard disc faster than floppy.

Control--are items weighted and measured?
    Yes, if patient is on check-back.

Medical Software Systems, Des Moines uses an optically read ticket--has a database.

Do you do inventory?
    No, it is done manually.

What institutions are using it?
    Some hospitals are awaiting installation.

Any assessment done by computer?
    No.

How do you report "holes" in diet, as a missing food?
    There's a space at end of each day for additions and comments.

When you first purchased your database from Case Western, did they give the programs or did you do it yourself?
    We did all the programs.

Do you not feel this is more costly as opposed to purchasing a program?
    There was not much available.

What kind of summaries are you printing from your program?
    Two page summary but a lot <u>incomplete</u>.

Do you have nutritional assessment on the computer?
    No.

COMMONALITIES AND DIFFERENCES
ARE WE ALL SPEAKING THE SAME LANGUAGE?

Grace J. Petot
Assistant Professor
Department of Nutrition
Case Western Reserve University

Each year this conference produces a new and longer listing of computerized nutrient data bases and systems. It is certainly not complete and perhaps never will be considering the fact that hardware of all kinds is very accessible; that software development has been greatly simplified; and that there is a burgeoning need for nutrition analysis in research, clinical applications, education, policy and program planning, and even the lay public. Nutrient data basic systems continue to proliferate in attempting to meet the needs. These activities are the basis for our common interests but we persist in speaking different languages as we separately pursue them.

Our contacts with users and developers of nutrient data base systems for a wide variety of applications have led to the production of a long list of knotty problems and questions. We, at Case Western Reserve University, attempt to deal with some; others we postpone in the hope that someone else will have a solution or answer sometime, perhaps at the next data bank conference. I would like to review some of them with you today. I do not have any solutions to propose, but bring them to you with the hope that together we can identify our commonalities and consider beginning to speak the same language.

DOCUMENTATION OF COMPUTERIZED NUTRIENT ANALYSIS IN PUBLISHED RESEARCH

Comparisons between and among dietary intake research studies is highly desirable and indeed is frequently done. Valiant attempts are made to qualify the similarities and differences, however usually not enough information is available in the methods section or in the bibliography to allow sufficient evaluation of the nutrient data used.

To illustrate some of these problems, I have reviewed all of the research reports in which dietary intakes were analyzed using a computerized nutrient data base system and which were published in the American Journal of Clinical Nutrition and the Journal of the American Dietetic Association from January 1982 through June 1983. These journals were selected because they contain numerous dietary intake reports which, when compared with other reports scattered among other research journals, appear to be representative. Summarizing the information presented a real challenge. Forty-two articles were identified in which computer calculations were made for 1-27 nutrients. Five of these reported research which was conducted outside of the United States. Twenty of the analyses were made on unidentified systems apparently developed at the institution where the research was done. These authors indicated the sources of nutrient values but the citations are inconsistent and, in some cases, vague, e.g. "from the literature", "from industry", "plus added values", "Handbook 8 extension", "from research and industry". Twenty-two authors used nutrient data base systems which were developed elsewhere. In only two of these reports was the date of the edition or version of the system provided or referenced. Table I summarizes the variety of food composition sources mentioned or cited in the bibliographies.

Often several sources are mentioned with a single report. However for only a few reports would it be possible to construct a duplicate data base using the references provided. Primary sources for most are the various forms of USDA Handbook 8. Twenty-two indicated that they added data from other sources to USDA tables or to the identified system. Only three authors, using unidentified systems, indicated that they used a machine readable version of Handbook 8 or 456. They were referenced as "Handbook 8-1963 tape". "Handbook 456-1977 tape" and "Handbook 456-3, Release 2, 1980". Each release of a USDA tape has a unique identifying number and year of release which does facilitate documentation.

Table I. Summary of sources of nutrient data cited in dietary intake studies published in the American Journal of Clinical Nutrition, Journal of the American Dietetics Association, January 1982-June 1983.

Sources mention or cited	Nutrient Data Base System: Unidentified n=20	Nutrient Data Base System: Identified n=22
Bulletin 72	3	0
Handbook 8, 1963	6	2
Handbook 456	10	3
Revised 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 two of these provided a date of the edition or version of the system.

A list of references will be supplied by the author on request.

For the other identified data base systems used, it might be expected that complete documentation of data sources would be provided by the developers. Thus this brief literature review illustrates that for purposes of comparability of dietary intake studies, identification of nutrient data sources is inconsistent at best and non-existent at worst. Editorial policies require scientific documentation for research methods and procedures but not for sources of nutrient information. Is it because nutrient data are unscientific or is it the absence of knowledge of the complexities of obtaining nutrient values for food?

USDA FOOD COMPOSITION TABLES

Printed food composition tables are becoming more expensive, less accessible, and less complete. Handbook 8 revisions are still not being widely used, especially in clinical applications. The total cost of a set of the revised books to date is approximately $70; by the time all revisions are complete, a set may cost well over $250 and will be unwieldy to use. Practitioners are not acquiring them for their personal convenience; therefore, are continuing to use Handbook 456 or accessing or purchasing data acquired, compiled and computerized into analysis systems by others.

It must also be recognized that purchase of a USDA tape requires the development of software for accessing, summarizing and reporting nutrient analyses. The most desired and most requested information for clinical practice and research is in the largest data set, USDA Nutrient Data Base for Standard Reference. This data set cannot be accommodated on a microcomputer and requires programming to obtain an easily readable, printed output. In addition, this data set is the most complete compilation containing updated information for foods in sections which have not yet been revised. These data may never again be widely available in a printed form at a reasonable cost.

Our work with the USDA Standard Reference Tape revealed a number of interesting facts. Although the food composition group at USDA provides an explanation of what is on the tape, it does not provide complete information about how it differs from the previous releases. I received some surprises such as the presence of imputed values in many of the unrevised food groups. Nutrients published in the Provisional Tables produced between tape releases are generally not on the tape and a few nutrient values for food in unrevised food groups have been changed.

INDUSTRY REPORTING OF FOOD COMPOSITION DATA

The food industry is generally quite cooperative in supplying nutrition information if it is available. However in trying to satisfy the diverse requirements of consumers' government regulations, dietitians and data bank developers, a variety of reporting formats appear. Consequently nutrient information may be received per portion, per 100 grams or per some other portion unit. Nutrient content may have been determined in the laboratory or calculated using unnamed versions of Handbook 8 or a combination of both and expressed as either absolute nutrient values or as per cent of U.S. RDA. As most data bases are constructed with nutrients per 100 grams, scaling of per portion values and translation of per cent U.S. RDA are then required. In addition to the opportunities for errors, we might question the practice of merging these data with the more precise values as expressed per 100 gram portion.

"Do you have brand name foods in your data base?" is one of the most frequently asked question of us. Yes we do, but it is not complete and completeness varies among the products. When we wish to examine dietary intake records comprised of both brand-name and conventional foods for a set of specific nutrients, do we fill in the missing values for the brand-name products? If so, what is the basis for the decision? Alternatively a conventional food or combination of foods with known nutrient values could be selected to replace the brand-name product. Again, how are the decisions made? These procedures also contribute to the problems of comparability between research studies.

Content of several nutrients and food components of current interest are defined and reported in a variety of ways. Among them are carbohydrates, starches, sugars, added sugars, refined sugars, dietary fiber, forms of Vitamin E activity and forms of Vitamin A activity. Members of the research community who design and review dietary intake studies define them their ways, the food industry reports them in a variety of ways and those of us who must reconcile them remain on the horns of many dilemmas. Again, we are speaking different languages.

FOOD NAMES AND CODING

Naming and describing of foods in the American food supply is a dynamic process; it has become even more dynamic with the growing trend to compress data bases into even smaller spaces in computers. We can all agree that the better the description of the food item, the greater the validity of the nutrient calculations. This problem is a joint one between the data gatherer, or the menu or recipe writer and the nutrient analysis system. Because we cannot expect everyone involved to be food scientists or food technologists, we must recognize that there are many levels of qualification of attributes of food and use them with corresponding appropriate levels of quality of data entered into the system for computation. The variability of food coding decisions is one of the most important factors which interferes with comparability of analyses from different systems.

NUTRIENT ANALYSIS OF RECIPES

An increasing use of nutrient data base systems is for analysis of recipes for the home, institutional food service or commercially prepared food service products. Standardized recipes for which all ingredients are precisely described, all procedures are controlled and for which all intermediate and final yields are determined will provide the most accurate calculated nutrient analysis. These conditions usually prevail for commercially produced products, much less frequently for institutional food production, and rarely in the home. We have found that nutrient calculations for some commercially prepared products have correlated very well with laboratory analysis. However, very close cooperation with the industry personnel was required to identify precisely all ingredients and yields. Of critical importance is the determination of cooked yields for the separate ingredients to obtain the best nutrient values for the finished product. USDA Handbook 102 is at present our best source of this information; however, it is far from complete.

SPEAKING THE SAME LANGUAGE

A review of past data bank conference proceedings provides evidence that we are speaking the same language as we identify common concerns and issues. During my brief review of only a few concerns, I am sure that many of you have thought of others. The rate of growth of nutrient data base applications is far greater than the speed with which we are seeking resolutions. We have a good start; now we need to speed up the process.

References Cited
Table I

Jones, B.R., E. Barrett-Connor, M.H. Critique, and M.J. Holdbrook, 1982. A community study of calorie and nutrient intake in drinkers and non-drinkers of alcohol. Am. J. Clin. Nutrient. 35:135-139.

Farris, R.P., M.S. Hyg, G.C. Frank, L.S. Webber, S.R. Srinivasan and G.S. Berenson, 1982. Influence of milk source on serum lipids and lipoproteins during the first year of life, Bogalusa Heart Study. Am. J. Clin. Nutrient. 35:42-49.

Dwyer, J.T., W.H. Dietz, E.M. Andrews, and R.M. Suskind, 1982. Status of vegetarian children. Am. J. Clin. Nutrient. 35:204-216.

Kerr, G.R., E.S. Lee, M. Lam, R.J. Lorimor, E. Randall, R.N. Forthofer, M.D. Davis, and S.M. Magnetti, 1982. Relationships between dietary and biochemical measures of nutritional status in HANES I data. Am. J. Clin. Nutrient. 35:294-308.

Roaij, J.M.A. Van, M.B. Katan, C.E. West, and J.G.A.J. Hautvast, 1982. Influence of diets containing casein, soy isolate, and soy concentrate on serum cholesterol and lipoproteins in middle-aged volunteers. Am. J. Clin. Nutrient. 35:925-934.

Danford, D.E., J.C. Smith, and A.M. Huber, 1982. Pica and mineral status in the mentally retarded. Am. J. Clin. Nutrient. 35:958-967.

Freeland Graves, J.H., B.J. Friedman, W. Han, R.L. Shorey, and R. Young, 1982. Zinc supplementation on plasma high-density lipoprotein cholesterol and zinc. Am. J. Clin. Nutrient. 35:988-992.

Stein, E.A., J. Shapero, C. McNerney, C.J. Glueck, T. Tracy, and P. Gartside, 1982. Changes in plasma lipid and lipoprotein fractions after alteration in dietary cholesterol, polyunsaturated, saturated, and total fat in free-living normal and hypercholesterolemic children. Am. J. Clin. Nutrient. 35: 1375-1390.

Faust, H., 1982. Effects of drinking water and total sodium intake on blood pressure. Am. J. Clin. Nutrient. 35:1459-1467

Buzzard, M., M.R. McRoberts, D.L. Driscoll, and J. Bowering, 1982. Effect of dietary eggs and ascorbic acid on plasma lipid and lipoprotein cholesterol levels in healthy young men. Am. J. Clin. Nutrient. 36:94-105

Smith, M.P., J. Mendez, M. Druckenmiller, and P.M. Kris-Etherton, 1982. Exercise intensity, dietary intake, and high-density lipoprotein cholesterol in young female competitive swimmers. Am. J. Clin. Nutrient. 36:251-255.

Garry, P.J., J.S. Goodwin, W.C. Hunt, E.M. Hooper, and A.G. Leonard, 1982. Nutritional status in a healthy elderly population: dietary and supplemental intakes. Am. J. Clin. Nutrient. 36:319-331.

Posner, B.M., C.L. Borman, J.L. Morgan, W.S. Borden, and J. C. Ohls, 1982. The validity of a telephone-administered 24-hour dietary recall methodology. Am. J. Clin. Nutrient. 36:546-553.

McGee, D., G. Rhoads, J. Hankin, K. Yano, and J. Tillotson, 1982. Within-person variability of nutrient intake in a group of Hawaiian men of Japanese ancestry. Am. J. Clin. Nutrient. 36:657-663.

Jain, M.G., L. Harrison, G.R. Howe, and A.B. Miller, 1982. Evaluation of a self-administered dietary questionnaire for use in a cohort study. Am. J. Clin. Nutrient. 36:931-935.

Hak, L.J., M.S. Leffell, R.W. Lamanna, K.M. Teasley, C.H. Bazzarre, and W.D. Mattern, 1982. Reversal of skin test energy during maintenance hemodialysis by protein and calorie supplementation. Am. J. Clin. Nutrient. 36:1089-1092.

Picone, T.A., L.H. Allen, M.M. Schramm, and P.N. Olsen, 1982. Pregnancy outcome in North American women. 1. Effects of diet, cigarette smoking, and psychological stress on maternal weight gain. Am. J. Clin. Nutrient. 36:1205-1213.

Omdahl, J.L., P.J. Garry, L.A. Hunsaker, W.C. Hunt, and J.S. Goodwin, 1982. Nutritional status in a healthy elderly population: Vit. D. Am. J. Clin. Nutrient. 36:1225-1233.

Gibson, R.S., B.M. Anderson, and C.A. Scythes, 1983. Regional differences in hair zinc concentrations: a possible effect of water hardness. Am. J. Clin. Nutrient. 37:37-42.

Todd, K.S., M. Hudes, and D.H. Calloway, 1983. Food intake measurement: problems and approaches. Am. J. Clin. Nutrient. 37:139-146.

Baecke, J.A.H., W.A. Van Stoveren, and J. Burema, 1983. Food consumption, habitual physical activity and body fatness in young Dutch adults. Am. J. Clin. Nutrient. 37:278-286.

Kromhout, D., 1983. Changes in energy and macronutrients in 871 middle-aged men during 10 years of follow-up (the Zutphen Study). Am. J. Clin. Nutrient. 37:287-294.

Stuff, J.E., C. Garza, Smith E. O'Brian, B.L. Nicholas, and C.M. Montandon, 1983. A comparison of dietary methods in nutritional studies. Am. J. Clin. Nutrient. 37:300-306.

Hambidge, K.M., N.F. Krebs, M.A. Jacobs, A. Favier, Sci D. Pharm, L. Guyette, and D.N. Like, 1983. Zinc nutritional status during pregnancy: a longitudinal study. Am. J. Clin. Nutrient. 37:429-442.

Hunt, I.F., N.J. Murphy, A.E. Cleaver, B.Faraji, M.E. Swendseid, A.H. Coulson, V.A. Clark, N. Laine, C.A. Davis, and J.C. Smith Jr., 1983. Zinc supplementation during pregnancy: zinc concentration of serum and hair from low income women of Mexican descent. Am. J. Clin. Nutrient. 37:572-582.

Beaton, G.H., J. Milner, V. McGuire, T.E. Feather, and J.A. Little, 1983. Source of variance in 24-hour dietary recall data: implications for nutrition study design and interpretation. Carbohydrate sources, vitamins and minerals. Am. J. Clin. Nutrient. 37:986-995.

Dewey, K.G., 1983. Nutrition survey in Tabasco, Mexico: nutritional status of preschool children. Am. J. Clin. Nutrient. 37:1010-1019.

Yeung, D.L., J. Hall, M. Leuna, and M.D. Pennell, 1982. Sodium intakes of infants from 1 to 18 months of age. J. Am. Diet. Association. 80:242-244.

Van Staveren, W., J.G.A.J. Hautvast, M.B. Katan, M.A.J. Van Montfort, and H.G.C. Van O'osten-Van Der Goes, 1982. Dietary fiber consumption in an adult Dutch population. J. Am. Diet. Association. 80:324-330.

Johnson, C.C., M.H. Stone, A. Lopez-S., J.A. Hebert, L.T. Kilgore, and R.J. Byrd, 1982. Diet and exercise in middle-aged men. J. Am. Diet. Association. 81:695-701.

Salz, K.M., N.Z. Haigh, G.A. Chase, and P.O. Kwiterovich, 1982. Fat and cholesterol intakes of white adults in Columbia, Maryland. J. Am. Diet. Association. 81:541-546.

Lillian, L.J., A.M. Huber, and M.M. Rajala, 1982. Diet and ethanol intake during pregnancy. J. Am. Diet. Association. 81:252-257.

Sempos, C.T., N.E. Johnson, P.J. Elmer, J.K. Allington, and M.E. Matthews, 1982. A dietary survey of 14 Wisconsin nursing homes. J. Am. Diet. Association. 81:35-40.

Marshall, M.W., and J.T. Judd, 1982. Calculated vs. analyzed composition of four modified fat diets. J. Am. Diet. Association. 80:537-549.

Brown, E.K., E.A. Settle, and A.M. Van Rij, 1982. Food intake patterns of gastric bypass patients. J. Am. Diet. Association. 80:437-443.

Sawicki, M., and J. Endres, 1983. Energy and nutrient calculations using an Optical Character Reader System. J. Am. Diet. Association. 82:135-141.

Hutton, C.W., and R.B. Hayes-Davis, 1983. Assessment of the zinc nutritional status of selected elderly subjects. J. Am. Diet. Association. 82:148-153.

Domer, J.A., 1983. Nutrition in a private day care center. J. Am. Diet. Association. 82:290-293.

Carter, P., D. Carr, J. Van Eys, I. Ramirez, D. Coody, and G. Taylor, 1983. Energy and nutrient intake of children with cancer. J. Am. Diet. Association. 82:610-615.

Short, S.H., and W.R. Short, 1983. Four-year study of university athletes' dietary intake. J. Am. Diet. Association. 82:632-645.

O'Hanlon, P., M.B. Kohrs, E. Hilderbrand, and J. Nordstrom, 1983. Socioeconomic factors and dietary intake of elderly Missourians. J. Am. Diet. Association. 82:646-653.

Khan, M.A., 1983. Sodium intake from meals and snacks consumed by college students. J. Am. Diet. Association. 82:664-666.

REPORT OF AN INTERNATIONAL MEETING: INFOODS INTERNATIONAL NETWORK OF FOOD DATA SYSTEMS

Carol T. Windham, Ph.D.
R. Gaurth Hansen, Ph.D.
Department of Nutrition and Food Sciences
Utah State University

Five months ago, during January 30 to February 5, 1983, a significant international planning conference was conducted to discuss the need for and development of a world-wide food data networking organization, to be called INFOODS, for International Network of Food Data Systems. The conference, which was held at the Rockefeller Conference and Study Center in Bellagio, Italy included participants from government, industry, academia and nutrition, health and agricultural institutes in a number of countries in North and South America and western Europe. Co-chairmen of the meeting were R. Gaurth Hansen, Utah State University and Vernon R. Young, Massachusetts Institute of Technology.

The immediate objectives of the INFOODS planning conference were to explore and develop relevant topic areas and to define needs and strategies leading to the establishment of a standardized, high-quality, readily accessible food data bank systems network. In order to meet this challenge, a series of pre-determined, but not necessarily inclusive, issues were raised for consideration. Among the issues addressed was the need for world-wide information on the chemical composition of foods and the desirability and feasibility of meeting this need on an international level. Also discussed were the difficult problems that must be faced in generation of quality food composition information; the development of a structure for standardizing data entries, records and files; and a method of systemitizing them in a way which will facilitate the user without exceeding the limits of current technology.

USER NEEDS

Food composition data are of value to many public and private agencies and individuals, including governments, food industries, research and educational institutions, physicians, dietitians, and increasingly, at least in the United States, the food-consuming public. The needs of these different users with respect to both data base content and form are necessarily varied.

A few of the uses identified by conference participants include:

- Performing nutritional assessments, diet evaluation and planning for normal populations and those with special needs
- Developing dietary standards of reasonable and adequate nutrient intake
- Identifying food and nutrient consumption patterns of population groups, evaluating the adequacy of consumption and trends over time
- Identifying relationships between food and nutrient consumption practices and the incidence of degenerative diseases such as heart disease and cancer

REPORT OF AN INTERNATIONAL MEETING:
INFOODS
INTERNATIONAL NETWORK OF FOOD DATA SYSTEMS

Carol T. Windham, Ph.D.
R. Gaurth Hansen, Ph.D.
Department of Nutrition and Food Sciences
Utah State University

Five months ago, during January 30 to February 5, 1983, a significant international planning conference was conducted to discuss the need for and development of a world-wide food data networking organization, to be called INFOODS, for International Network of Food Data Systems. The conference, which was held at the Rockefeller Conference and Study Center in Bellagio, Italy included participants from government, industry, academia and nutrition, health and agricultural institutes in a number of countries in North and South America and western Europe. Co-chairmen of the meeting were R. Gaurth Hansen, Utah State University and Vernon R. Young, Massachusetts Institute of Technology.

The immediate objectives of the INFOODS planning conference were to explore and develop relevant topic areas and to define needs and strategies leading to the establishment of a standardized, high-quality, readily accessible food data bank systems network. In order to meet this challenge, a series of pre-determined, but not necessarily inclusive, issues were raised for consideration. Among the issues addressed was the need for world-wide information on the chemical composition of foods and the desirability and feasibility of meeting this need on an international level. Also discussed were the difficult problems that must be faced in generation of quality food composition information; the development of a structure for standardizing data entries, records and files; and a method of systemitizing them in a way which will facilitate the user without exceeding the limits of current technology.

USER NEEDS

Food composition data are of value to many public and private agencies and individuals, including governments, food industries, research and educational institutions, physicians, dietitians, and increasingly, at least in the United States, the food-consuming public. The needs of these different users with respect to both data base content and form are necessarily varied.

A few of the uses identified by conference participants include:

- Performing nutritional assessments, diet evaluation and planning for normal populations and those with special needs

- Developing dietary standards of reasonable and adequate nutrient intake

- Identifying food and nutrient consumption patterns of population groups, evaluating the adequacy of consumption and trends over time

- Identifying relationships between food and nutrient consumption practices and the incidence of degenerative diseases such as heart disease and cancer

- Developing dietary guidance tools and education programs for individuals and populations

- Facilitation of the food and agricultural industries in their efforts to provide a nutritious, safe and acceptable food supply

- Facilitating governmental regulation of the food supply and maintenance of nutritional quality

- Establishing standards of identity for traditional foods, requirements for imitation foods and standards for labeling and advertising claims

- Determining interactions between nutrients; and toxicological, pharmacological and other aspects of the diet

- Assessing changes in the nutrient/contaminant/substance content of the food supply

Given the variety of needs and users and major differences in foods and their preparation in different regions of the world, it was essential to address the question of whether or not it is feasible or even desirable to establish an international food data bank system. As a result of the INFOOD planning conference it is possible to say yes to the second question--it is desirable--and, given commitment and participation by the international scientific community, it may even be feasible.

WHAT INFOODS HAS TO OFFER

The first major step is to define what INFOODS is to be. Many countries have their own national food and nutrient data bases. There are also a number of public and private data bases housed in various agencies and institutes throughout the world. Although these systems may not be meeting all the needs of their respective users, generally efforts are being made to update and revise data to keep pace with new analytical techniques, with professional interest in an increasing number of nutrients, contaminants and other food components, and with the continually changing nature of food supplies. It is unrealistic to expect in the near future that existing data bases will be modified and merged to form one large single international data base.

The question then becomes what can INFOODS--both as an international data base system and as an organization--offer to potential users? What can INFOODS contribute over and above what is already being provided by public and private data bases to those who need and use them?

One of the primary contributions can be to provide direction and assistance to those countries and regions of the world currently without data base systems. This is particularly important in developing areas of the world where the resolution of food and nutritional problems is impeded by the scarcity of reliable data on the nutritional value of local foods. The food composition data which are available in Latin America, for example, are scattered among a number of tables prepared in different countries and much of the information is unpublished. There is also a lack of uniformity among tables in the way in which data have been obtained, organized and presented.

The Food Composition Table for Use in Latin American was compiled and published by INCAP-ICNND in 1961 to "meet an urgent need for expressing all available analytical data in a uniform manner and assembling them into a single table in which foods could be identified by their scientific name or by any one of their popular names." These data have not been updated in the 20 years since they were published and do not necessarily represent foods currently being consumed. Similar problems exist in other developing regions of the world.

At present there is no internationally organized effort to develop a standardized system to collect and process food data for common use in a developing regions, although the data would be valuable and used if available. INFOODS has the opportunity in these regions to provide leadership in formulating international standards and guidelines for the development of new food data systems in order to assure compatibility among new systems and facilitate data exchange.

Another area where INFOODS can play a unique role is in the provision of world-wide accessible data on foods commodities used in international trade and commerce. One participant at the INFOODS conference indicated that his country must import wheat from a number of other countries in the world. As there is a concern regarding selenium nutriture in his country, he is obliged to sample each boat load of wheat at the docks and analyze it in his laboratory for selenium, so as to have a measure of levels of this nutrient in the food supply. He may well have the best data in the world on selenium content of different varieties of wheat produced in different regions of the world. But how much more useful and convenient it would be for him and others to have access to the data from a computer terminal, either for determining nutrient composition after importation or for use in decision-making before finalizing trade agreements regarding food commodities. Nutrient composition data on foods of international commerce would be invaluable to those countries for which food security is a national concern, because of their vulnerable trade positions and dependence upon food imports to meet the basic energy needs of their populations.

Undoubtedly, one of the most immediate and critical needs in the realm of food nutrient data bases today, is that of standardization. Currently, data bases are internally consistent at best. Because of differences in nomenclature and in kinds and number of foods and food components included, it is difficult, if not impossible, to use more than one data base interchangeably or for supplementary or verification purposes. These problems are compounded by lack of documentation to indicate how or where the data were obtained or how reliable the data entries are. Without some type of standardization, documentation or guidelines, it can only be expected that food composition data bases will become less and less compatible, thus further limiting the use of food component data as a resource for international cooperation and investigation. Guidelines are needed for the generation, compilation and reporting of food component data; for data quality indices based upon sampling procedures, analytical methodologies and lab practices; for criteria for accepting data from laboratories or the literature and for imputing values; and for uniquely naming and classifying foods and food components.

TASKS FOR INFOODS

The consensus of the INFOODS conference group was that information on the composition of foods throughout the world contributes significantly to a variety of national and international activities. However, these activities are being hampered by:

- the limited amount of current, reliable, easily accessible data on the composition of the foods of the world

- the lack of compatibility among the existing food composition data bases

- the lack of national/international standards and guidelines for gathering, storing and disseminating food composition data

- the amount of food composition data being gathered by inadequate and conflicting techniques

In order to address these issues the conference working groups made a number of recommendations as to tasks that should be undertaken by INFOODS.

The working group concerned with user needs suggested that INFOODS be responsible for organization and coordination of an expert group to detail the specific needs of the communities of users for food composition data. This expert group should include representatives from the international community and involve participation by the food industry and governmental administrations. Another suggested task was organization and coordination of an expert group to consider what specific data bases and subsets of worldwide food composition data now exist or would be of value to the mission of INFOODS. Here special attention should be paid to problems unique to developing regions in the world.

The working group which considered data base content, suggested that INFOODS be responsible for formation of an expert group to examine methods and criteria and to develop guidelines for extracting archival data from the literature and alternative sources. This expert group should consider the critically important question of criteria for accepting data into any food component data base. Also suggested by the working group was the implementation of a global survey of existing data bases and banks of ongoing and future collection efforts.

A third working group considered the question of how to gather necessary data, insuring its accuracy and completeness. They suggested that INFOODS organize and coordinate an expert group to explore the issue of quality of data and, specifically, to update current guidelines for the preparation of tables of food composition. It was recommended that an updated publication of such guidelines should cover at least: (a) generation of food composition data by analysis (direct method), including sampling methods, food laboratory practice (reference food material, quality control, procedures, modes of expression of data and standardization of conversion factors, and approaches for data appraisal and interpretation; (b) use of food composition data from other sources (indirect method) involving using data from the original literature and calculating and imputing analytical values and (c) training of analysts for food composition data generation.

This working group also suggested selection of an expert who could devote 6-9 months on a full-time basis to researching and preparing an updated text of the proposed guidelines. This individual would undertake visits to laboratories in developing and industrialized countries prior to the period devoted to full-time text preparation. Another suggestion was that INFOODS investigate the establishment of a program of training fellowships to enable the broader development of skill and expertise in all areas directly concerned with food composition. Specifically, mechanisms should be developed by INFOODS, through international agencies and other suitable organizations, to facilitate the implementation of this fellowship program.

Also recommended was the investigation of the feasibility of establishing an international journal devoted to food composition studies. It was proposed that a journal would facilitate adoption of guidelines by the scientific community, serve as an information source for any future revision of the guidelines, and would provide a means for dissemination of findings and critical reviews in all areas of food composition research.

A fourth working group was asked to consider data base organization and content with respect to the diversity of data required and the problems which arise in managing these data. These are guidelines which could be used by new, developing data bases if desirable. This group suggested that INFOODS organize and coordinate an expert group to establish nomenclature and a system of coding to be used in INFOODS. The expert group should plan, define and recommend terms for: identifying food and food components, units of expression, analytical methods, preferences, locations, environmental conditions, and others. Membership of this group should include individuals with knowledge of foods in international trade and raw and processed foods in developed and developing countries.

Also suggested was that INFOODS organize and coordinate an expert group to explore and plan the informational networking system aspects of INFOODS. This group should (a) develop a system concept in terms of data flow, data regulation and information services to be provided; (b) discuss with existing centers those prepared to serve as INFOODS regional centers with a view to determining how they are to be integrated into an overall system concept and how their current structure and modes of operation influence the networking concept; (c) develop a planned implementation plan. A third recommendation of this working group was that INFOODS organize and coordinate an expert group to develop information interchange standards for food composition data. This group should develop a standard format and set of conventions for the interchange of food data between regional centers and, if desired, between laboratory and regional centers. The format arrived at should be usable in communicating with both large and small systems and should be designed independently of internal formats of any particular machine.

IMPLEMENTATION AND MANAGEMENT

In order to implement the recommendations made by the various working groups, a fifth working group explored the organizational framework necessary to plan and develop the structure and functions of INFOODS. It became obvious at the conference that INFOODS needs to have several very different aspects. It needs to be (1) a network of regional data centers; (2) an organizational and administrative framework for various expert task forces; (3) the generator (commissioner) of special international data bases; (4) a general and specific resource for persons and organizations interested in food composition data on a worldwide basis.

The conference group proposed that such an organization be set up to assist in the improvement of the state of food composition data around the world. This organization, called INFOODS, for International Network of Food Data Systems, would establish an international network of food composition data bases, identify all possible additional sources of food composition data, and direct and coordinate the development of guidelines and standards for the collection, storage and interchange of food composition data.

MISSION

The mission of INFOODS is to promote international participation and cooperation in the acquisition and dissemination of complete and accurate data on the nutrient composition of foods, beverages and their ingredients, in forms appropriate to meet the needs of government agencies, nutrition scientists, health and agriculture professionals, policy makers and planners, food producers, processors and retailers, and consumers.

ORGANIZATION

It was agreed by the participants that the INFOODS organization would consist of a secretariat (Dr. William Rand, executive secretary) guided by an executive committee and advised by a policy committee (Dr. Vernon Young, chairman of both), to be headquartered at the Massachusetts Institute of Technology. This executive committee will help to set up regional liaison committees in specific geographic areas.

The organization at this point has three major functions: (1) to serve as an active clearing house for information on all aspects of food composition data; (2) to establish an international network to facilitate communication between existing data bases; (3) to sponsor task forces needed to carry out its mission of improving the state of food composition data.

POST-CONFERENCE ACTIVITIES AND PLANS

The United Nations University (UNU) has agreed to serve as the sponsoring agency and INFOODS has been approved by the UNU council as long term project of its Food, Nutrition and Poverty Program. Close working relationships are being established and FAO and WHO. FAO has a continuing responsibility for the development and distribution of regional food tables and INFOODS has a mandate to cooperate with FAO in this task as desired by them. Similarly, FAO and WHO have responsibilitY for the Codex Alimentarius and INFOODS will not initiate any activities that would infringe on it but will cooperate with FAO and WHO as they consider appropriate.

Dr. William Rand met with a European data base group in Amsterdam at the end of May and plans to meet with Latin American representatives in Miami in August 1983, and with Asian representatives in November, 1983 to organize regional committees to aid and assist INFOODS.

A regional liaison committee has been set up for Northern America with the proposed responsibilities of assisting INFOODS by (1) gathering of data within the region, (2) identifying personnel and technical resources, (3) identifying special problems, needs and resources, (4) helping to obtain funding, and (5) contributing to policy decisions with its chairman serving on the INFOODS policy committee. Dr. Alex Campbell, Food and Nutrition consultant in Ottawa Canada is the chairman and Dr. Frank Hepburn from USDA Human Nutrition Information Service is the secretary of this regional committee.

Most of the other activity since the planning conference has been directed toward exploring and formalizing the recommendations that were made by participants and working groups. As a result of these recommendations the following activities have been planned:

- A survey of existing data bases to lead to a directory of international food composition data bases that will be the forerunner of the network. INFOODS will be able to provide information on the existence, quality, accessibility of specific food composition data around the world. This will be an ongoing activity, and the first directory is expected within a year.

- A data quality task force, to develop criteria and guidelines for the quality of entries in food composition data bases. This effort will include examination of sampling, analytical methodologies, good laboratory practices, modes of expression of data and conversion factors, as well as criteria for accepting data from the literature and for calculating and deriving data from analytical values. A task force has been organized for this effort with a report expected within a year. Dr. David Southgate of the British Agricultural Research Council is the chairman.

- A classification and nomenclature task force to examine the ?problem of uniquely characterizing food and food components to include !a) a review and evaluation of existing nomenclature systems and methods of classification and descriptions, and (b) the recommendation of international standards.

- A data base form and content task force to develop international guidelines for the content and form of an "ideal" data file, to include (a) items and components to be included, (b) additional information needed in the data file, and (c) logical formats of data records and files for storing, transferring, and distribution of food composition data.

- An information systems task force to advise on the design and to monitor the information systems aspects of INFOODS and the regional FOODS; to consider initially (a) the constraints that the data bases, users, and gatherers of the data put on the system and (b) the limitations that technology and being part of a network put on the individual components. Dr. Wolfgang Trebejahr of WHO is the chairman.

- A task force to survey the users of, and needs for, food composition data to include (a) magnitude and types of users and how frequently they use the data, (b) what data are most frequently used, requested or needed, and for what purposes, (c) what unmet needs for food composition data exist, and why these needs are unmet, and (d) what sort of network could best serve the users.

- Setting up and coordinating additional task forces as recommended by the policy committee as being essential for the mission of INFOODS.

- Organization of regional liaison committees for support of INFOODS by identifying special problems, needs and resources of specific geographic regions. It is hoped there will be liaison committees to represent the interests of South American, Europe, Asia, Africa and others.

- Investigation of the feasibility of establishing an international journal with the working title, "INFOODS - An International Journal of Food Analysis, Food Composition and Food Data Systems," that would provide a focus for the publication of research and the dissemination of information in the broad field concerned with the generation, processing, and use of food consumption data.

The success of INFOODS and the contribution it will make depends upon the extent to which you, the participants in the 8th Annual Nutrient Data Base Conference, and others in the scientific community contribution to its development. INFOODS welcomes any suggestions, expertise, time and other resources that you have to offer.

(The authors thank Dr. William Rand for useful discussions and contributions to this presentation.)

ASTM--AN INTRODUCTION

John J. Rothrock and Kenneth C. Pearson
American Society for Testing and Materials

I appreciate this opportunity to discuss the American Society for Testing and Materials (ASTM), what it is, what it does, how it does it. First, however, I would like to dispel some myths about ASTM. Many people think that ASTM is a government-affiliated organization. It isn't. Some think that ASTM is a lobbying organization. It isn't. Others charge that ASTM is an industry-dominated organization. ASTM isn't dominated by any interest group. What, then, is ASTM?

ASTM is a non-profit corporation organized in 1898 for one purpose and one purpose only...the development of full voluntary consensus standards...with the emphasis on full consensus. That simply means that in all of our standards work, everyone who has an interest in a standard may have a say in its development.

ASTM is a management system for the development of standards for materials chiefly known for products, systems, and services. It provides a legal, administrative, and publications forum within which all of the interested parties can meet on a common ground to write standards which will best meet the needs of all concerned.

I somehow believe your concept of "Common Conventions" embraces our view of organized standardization. ASTM defines standardization as: "The process of formulating and applying rules for an orderly approach to a specific activity for the benefit and with the cooperation of all concerned."

In the ASTM process, all points of view come to the standards table... producers, users, ultimate consumers, and the general interest representatives of government and academia...all with their own points of view, their own biases clearly labelled. It's frequently an adversary environment and, thus, can be a time-consuming process. We do not believe in unbiased experts...all of the differences and biases are brought into the ASTM forum, presented at the standards meetings, and then eventually consensus is developed. But we truly believe that the standards produced will have the highest credibility and will meet the needs of all concerned, because all concerned have had an opportunity to have their voices heard.

Quite often, people think some magical body of expertise exists in ASTM that writes and develops all of the standard documents. ASTM headquarters has no technical research or testing facilities. Such work is done voluntarily by those who work within the ASTM system-technically qualified members of the society located throughout the world. ASTM headquarters doesn't, however, have the organizational experience and skills to make this work fruitful and rewarding.

ASTM ORGANIZATION

Management of the society is vested in a board of directors. Board committees cover:

- Finance
- Society development (biennial updating LPR, membership, new developmental activities, public relations, promotion)
- Technical committee activities
- Overall voluntary standards system (role of ASTM interface private/governmental and national/international)

Separate standing committees of the society--related more directly to the great universe of the ASTM technical committees ("acronyms are awful").

- COTCO (regulations & committee operations)
- COS (oversee procedural requirements and due process)
- COP
- COT (uniformity and consistency--difficult--liaison)
- Research & technical planning

The core of ASTM is technical committees (TC's). We have a great diversity and encompass a very broad umbrella.

Some of the more common or well-known technical committees include:

- Steel
- Petroleum
- Cement & concrete
- Non-ferrous metals
- Paint
- Nuclear
- Solar
- Space technology
- Water
- Air/atmospheres pollution
- Environment/biological effects
- Resource recovery

Some of the less familiar committees include:

- Sensory evaluation
- Food service equipment
- Meat and poultry
- Forensic sciences
- Fire standards
- Computerized systems

And some of the more surprising include:

- Medical/surgical devices
- Orthotics
- Sports equipment (safety)
- Football helmets
- Skiing
- Amusement rides/devices
- Consumer products (high chair, lighters, tub/shower devices)

The umbrella is unlimited--those involved determine the need for coming into the ASTM management system for standardization.

Policy, of course, set by board, but hierarchically somewhat the reverse of industry. Most organizations inputs for policy and operations come from the TCs to ASTM headquarters for the standing committees consideration, and then to the board. Much less absolute direction from the "top-down" so common in industry!

Currently, ASTM has more than 31,000 members including some 3,000 international members. Since many members serve on more than one committee, total unit participation is well over 80,000.

Being non-profit, how is ASTM financed? What is the cost of writing standards in ASTM? First of all, ASTM does not charge for its services; there are no project costs.

Annual budget - $14M. Approximately 15% of ASTM's money comes from return on investment and administrative fees ($50 a year for society membership--belong to any number of technical committees).

Most of this money is returned to the members, however, in the form of a free part of the annual book of ASTM standards, reduced rates on additional parts and on other publications, a subscription to our monthly magazine Standardization News, and an enormous amount of administrative back-up which is provided by the headquarters staff.

The other 85% of ASTM's money comes from the sale of its publications.

- 66--volume annual book of ASTM standards containing over 7,000 standards under copyright by ASTM
- Separate--reprints of individual standards
- Special technical publications (STP's) which result from the 35-40 symposia ASTM committees sponsor each year.
- Journals and compilations
- Scientific data series

This financial picture represents just the "tip of the iceberg" in the costs of writing standards. It is estimated that 10-12 times the $14M that flows through ASTM headquarters is expended in the overall development of standards. Expenses are incurred, however, for travel and lodging in attending ASTM meetings, members' use of their own laboratories/facilities for testing and research, and most important...the member's time. But the organizations/companies they represents are the ones supporting these expenses.

TYPES OF STANDARDS

Commonly a noun--ASTM is also an adjective in titling ASTM documents

Standard (with designation) has gone through full consensus balloting procedures and met all requirements (proposals)

Six types of standards are listed with a brief/concise differentiation.

1. Standard specifications which define the boundaries or limits on the characteristics of a material, product, system, or service.

2. Standard methods of test which prescribe ways of making given measurements (thus supporting specifications).

3. Standard practices which suggest accepted procedures for performing given tasks.

4. Standard terminology/definitions which create a common language for a given area of knowledge.

5. Standard guides which propose an approach, a series of options or instructions that do not prescribe a definitive/specific course of action.

6. Standard classifications which set up categories in which objects or concepts may be grouped.

I believe efforts such as yours concerning a nutrient data base would make use of practices, guides, classifications and terminology.

A misconception about ASTM standards is that they are mandatory. This is probably based on their widespread use and acceptance throughout the country...as a matter of fact, all over the world. ASTM standards are developed voluntarily and they are used voluntarily. They become mandatory only when referenced by a regulatory agency, such as all levels of government building code authorities. This is not uncommon, but generally speaking, ASTM standards are voluntarily prepared and voluntarily used.

Now, I would like to describe in more depth, TC's and the standards development process.

Presently there are 140 standards-writing committees, about 1,800 subcommittees and thousands of sections and task groups.

Anyone who is knowledgeable in the work of a committee may serve on that committee. ASTM's bylaws preclude the exclusion of any qualified person who wishes to serve. And there is no restriction as to the number of persons who may participate on a committee. Some committees have as few as fifty, while others have a thousand or more. The average committee may have about 200-250 members.

The only restriction ASTM lays on its committees is that they be in balance. This simply means that the number of voting producers/manufacturers/vendors on committees having a commercial interest cannot be greater than the number of users, consumers, and general interest voters. Thus, the producers viewpoint cannot dominate the action when standards are being voted upon. These are known as classified committees. Committees that do not write specifications and have no commercial interest are not required to be classified and balanced. Such determinations are made by the committee itself, subject to approval by the board.

TECHNICAL COMMITTEE MAKE-UP

Main--Accumulated voting interests represented for balloting broad scope of activities.

Subcommittees--separate voting interests for balloting (split up), more definitive scope, specific tech activity, separate entities, specific technical expertise, membership responsible for their own meetings, agendas, minutes, reports, etc.

Technical subcommittees structured--material/product type, type of standards, sequence of testing and data, properties (whatever committee decides). Technical/service subcommittees support others--edit, statistics/data handling, terminology.

Executive--Executive direction/administration arm for overall committee. Overall same as main, plus subcommittee chairman, or members at large. Executive representative of overall committee organization administrative functions delegated to subcommittees/groups (as illustrated)--LRP, symposia, meetings, recognition/awards, liaison, international, research, education/training, etc.

Sections/task groups--(based on size and extent of study projects) There are relatively small groups with definitive assignments. They are responsible for the initial drafting and preliminary testing/screening for applicability. They draft a document for subcommittee, which is accompanied by data supporting standard (inter-lab testing, application, precision, research report)

TYPES OF ACTIONS

Listed below are the types of action the committees are involved in:

New standards (of course)--but more than new development

Revision--"Not cast in concrete"--updated, revised as often as necessary, kept abreast of technology/current practice

Reapproval--review yearly, but if no changes in 5 years, action to reapprove (exposure to all members), such actions indicated in designation of standard

Withdrawal--obsolete, replaced by other standard(s)

STANDARDS DEVELOPMENT & BALLOTING PROCESS

- All written letter ballots (mailed to all members of committee roster).
- Comments can accompany affirmative.
- Negative votes accompanied by written explanation.
- Not-persuasive--responding written explanation.
- Ballot tallies and documentation--permanent record file.
- Not uncommon--single negative persuasive, rework, reballot.

Each ballot iteration creates a greater exposure to more people with the goal to generate inputs and improve the document and make it more viable.

Consensus in ASTM is not unanimity

Appeal mechanisms are set up throughout the organization.

In all of this standards work and the overall committee operations, ASTM headquarters provides the technical committees with administrative support. An ASTM staff manager and administrative assistant is assigned to technical committees. They attend the committee meetings and are responsible to make certain that ASTM procedures are being carefully followed and to insure maximum access by the committee to ASTM support facilities. The goal is to take most of the administrative load off the committee members, thus allowing them maximum time to devote their efforts to the important job of writing standards. ASTM support services include preparation and mailing of ballots, tabulation and summary reports of main committee and society balloting, duplication and mailing of meeting notices/agendas/minutes, meetings arrangements and coordination including securing of meeting rooms and hotel accommodations, a phenomenal editing department (of course), publication of the standards produced by the technical committees, maintenance of membership records, membership recruitment, standards status records, publicity, promotion, etc.

What all of this really provides is continuity--follow-up, towards the ultimate goal of developing standards by those concerned and in need of standardization.

CONCLUSION

If you divide the responsibilities in the ASTM made of standardization, you, the standard writers/developers, are in charge of what is done and the society is somewhat in charge of how it is done--based on the rules and regulations.

You provide the technical expertise/activity and develop the needed type of standards. We help you organize and carry out and efficient, coordinated standards program--utilizing our expertise in management and administrative support.

ASTM has tremendous credibility--including both its published standards and its management skills.

It has been proven time and time again, that ASTM is a management system that lends itself to virtually any standards need.

My goal was to describe standardization, the ASTM way...and for you to determine whether the needs exist, and whether the interests and active participants are there to fulfill those needs.