Panel: Moving Into the Next Century: Databases

Loretta White Hoover, Ph.D., R.D., University of Missouri-Columbia

In the next century, computerized databases are likely to be the primary media for tables of food composition with printed tables replaced by alternative media such as optical disks, FAX, or videophone. Already, we have seen that comprehensive printed tables are bulky, expensive to produce and acquire, and difficult to keep up-to-date. Electronic data transfer offers the opportunity for selective retrieval of data of interest from depositories. Is a centralized clearinghouse for nutrient data in our future?

As we envision databases in the future, some of the aspects of databases to consider are their structure, contents, documentation, management, and software. Understanding the factors that are impacting each of these aspects of databases enhances our ability to foresee the characteristics of nutrient databases.

Nutrient Database Structure

From a structural standpoint, nutrient databases will probably have a complex configuration. Many of you will remember when a nutrient database was a deck of computer cards. With enhancements in computer technology, we have begun to separate data according to type of data such as nutrient values and food descriptions.

In addition to familiar files of nutrient values and food descriptions, other types of coordinated data, linked by relational keys, may become integral aspects of nutrient databases. Some of the related data files may include: nutrient descriptions, weights and measures, yield factors, nutrient bioavailability factors, nutrient retention factors, data quality, refuse factors, cross-references, code translation tables, ingredients, recipes/formulations, external link codes, food labels, and portion model images. A more complex configuration has implications for data management and software development. With data in the related files linked by relational keys, data redundancy can be minimized. Coordination of these collections of data will reduce reliance on printed documentation and will support on-line, real-time look-up of pertinent data in a software application.

At some point, we will probably agree that values are needed for only one standard quantity such as 100 grams in our databases. Coordinated records in a weights and measures file will provide the flexibility to process any conceivable portion reported on a dietary record. The computer will assume the role of determining the quantity of each nutrient, and we will have minimized the number of values that must be maintained for each food.

Nutrient Database Contents

The contents of nutrient databases will expand to include new data elements, new foods and modifications of traditional foods, new commercial ingredients, and additional nutrients and non-nutrients. In particular, new data elements will be created to store detailed documentation and pertinent dates, fully qualified with century indicators. Trend analyses, product tracking, and reprocessing of prior data create the need for a variety of dates such as: effective, introduced, expired, departure, or obsolete.

The expanding diversity in foods will result in the addition of many more foods and probably new groups of foods. Growing ethnic minorities and the rise in transnational agribusiness are acquainting us with an array of exotic and imported foods. Similarly, regional variations in food preparation require alternative versions of the same food. For example, chili as I know it from Texas is not the same dish that is served in other regions of the country.

The contents of nutrient databases will be influenced by consumer interest in diet and health and by the dietary requirements of individuals with chronic diseases or genetic disorders of metabolism. The number of foods will multiply as we include low-sodium, low-fat, fat-free, sugar-free, and gluten-free variations for many of our traditional foods.

Tracking the intake of individuals consuming an increasing proportion of their diet from manufactured foods requires that we conquer the brand name challenge. New ingredients such as fat replacers, fibers, gums, and protein components will yield nutrient profiles differing greatly from traditional forms of many foods. Incorporating data for individual products will probably cause us to rethink some of the groupings of foods in nutrient data bases. Also, we may seek ways to form composites of similar foods like granola bars or low-fat crackers. Unique identifiers, appended to existing systems for groupings, will facilitate organization of the overwhelming amount of data records reflecting the burgeoning food marketplace.

Mixed dishes, coded according to recipes, will probably constitute an increasing proportion of the data records in nutrient databases as we attempt to reflect data for foods as consumed. The recipe strategy provides a way to reflect ethnic and regional variations, to estimate values for constituents of interest when laboratory analyses are not feasible, and to recalculate nutrient profiles when ingredients change or the nutrient values of the ingredients are up-dated. Mathematical estimation of formulations for manufactured foods will probably be more prevalent as a technique for approximating values for nutrients not supplied by food processors. The expanding use of these calculation procedures will emphasize the essentiality of better data about cooking losses and gains and nutrient retention.

Nutrient databases will expand to accommodate more constituents linked to concerns about diet and health. Already, data about sugars, starches, alcohol, carotenoids, selenium and other minerals, specific forms of fatty acids, and some non-nutrient constituents are being added to nutrient databases. The amount of imputed values in nutrient databases will probably rise because data for these additional constituents will be available for only a limited number of foods. The problem of missing values will become more acute.

Nutrient Database Documentation

As the need for nutrient data continues to outpace the availability of analytical data, documentation of the types and quality of data will become more prevalent. For example, we will probably employ schemes for identifying if data were analyzed and by what method, normalized and by what basis, calculated and by what method, estimated and with what rationale, or imputed and by what method. In terms of data quality, related data records will provide detail about analytical methods, sampling plan, reference materials, and confidence indicators. Incorporation of documentation into nutrient databases will be facilitated by standard nomenclature and acronyms evolved by consensus.

Nutrient Database Management

Management of nutrient databases will become increasingly more demanding, time consuming, and expensive. Quality control protocols will be essential to insure the integrity of nutrient databases. Incoming data will be screened automatically by diagnostic routines, evaluated against previous data, and subjected to statistical tests. Judgmental actions will be documented and date stamped in the course of transaction processing to provide complete audit trails of data revisions.

Collaboration among the food industry, research organizations, and governmental agencies will facilitate compilation of truly representative databases.

Partnerships will probably be established to achieve a cost-effective strategy for compiling and distributing data for an ever-increasing number of foods and constituents. A clearing house, voluntarily supported by all players, would be a natural product of such coordinated efforts.

Multimedia technology will continue to offer new and different strategies for storage and dissemination of nutrient data. If developments in technology continue at the current pace, we can anticipate many more advances in computing and communications capability that will facilitate accessibility and management of nutrient databases.

Nutrient Database Software

Numerous software options will be standard to support query, data collection and coding, updating and maintenance, recalculation, iterative operations, and diagnostic evaluation. Software will be created or revised to accommodate new nutrient database designs reflecting new field sizes, new data fields, and new data files.

Innovations in software are expected as developers employ extensive use of graphics for portion models and results of analyses. Mathematical models will be used more extensively. Some examples of the use of mathematical models are linear programming for determining the proportions of ingredients in formulations, materials requirements planning (MRP) to determine food usage or requirements, and statistical models for testing outliers or forecasting data needs.

Multimedia technology such as optical disks, FAX, and videophone offer new possibilities for software development and enhancement. Data acquisition and communication will be facilitated by electronic data interchange and electronic file transfer.

Predictions About Databases

Even though my crystal ball is a bit hazy, I offer the following predictions about nutrient databases in the next century. In my opinion, the trends that we can anticipate relative to these aspects of nutrient databases are:

Structure	→	More Complexity
Contents	→	More Extensive
Documentation	→	More Specificity
Management	→	More Expensive
Software	→	More Capable.

In summary, nutrient databases will become even more integral to professional activities related to nutrition, diet, and health. The increased complexity of nutrient databases will require more sophisticated users. Assuring the desired contents in nutrient databases will require a strong sense of collaboration among data generators, compilers, and users.