DESIGNING A DATA BASE FOR HEALTH CARE APPLICATIONS

Diane M. Clapp, M.S., R.D.

The Ohio State University Hospitals

Introduction

The Ohio State University Hospitals is a 1,000 bed teaching hospital located on the Ohio State University campus in Columbus, Ohio. The hospital supports the needs of patients in the community and state through direct patient care, outpatient services and medical research. To provide for the nutritional well being of hospital patients, the need for extensive as well as reliable information on the nutritive value of foods is essential. The Ohio State University Hospitals' Department of Nutrition and Dietetics' Nutrient Data Base (NDB) was developed for in-house use for the therapeutic dietitian involved in dietary planning and patient education, as well as the medical research community.

In my presentation, I will be discussing the background of the OSUH NDB, considerations for nutrition software design, and our systems applications at OSU Hospitals.

Historical Background

The Ohio State University Hospitals recognized the need for an all-inclusive reference source for nutrient information as early as 1973 (Schaum, K.D., M. Mason and J. Sharp. Patient-oriented dietetic information system. J.A.D.A. 1973). As part of a Masters thesis, an extensive food composition table containing approximately 3600 food items was compiled and stored as a nutrient data bank on magnetic tape media. The file contained storage space to expand to a maximum of 9,999 food items, with 63 nutrient components for each item.

Format and data selection criteria for this study were as follows:

a) Nutrient values for 100 gram edible portions were to be stored in the data bank
b) No data published prior to 1960, with the exception of USDA Home Economics Research Report No. 4 were to be used
c) No foreign data were accepted. If more than one source contained data for the same food item, then:
    1. Data were selected from the study designed with better or best analytical methods or sampling procedures.
    2. The most recent data were chosen if analytical method and sampling procedures were comparable.
    3. Original data were selected rather than compiled data.

The USDA Agricultural Handbook No. 8 data cards were machine converted to our desired format. Form letters were sent to various food manufacturers whose products were used at University Hospitals and by individuals in the Ohio area. A MEDLARS (Medical Literature Analysis and Retrieval System) computer search was obtained which provided a bibliography of all the articles related to food analysis published in the journals listed in Index Medicus since 1964.

The original data bank took approximately two years to design, compile and store on magnetic tape. Although the file contained many gaps with regards to missing values, it was considered one of the most complete files available for that time.

OSUH Computer Center wrote programs for our mainframe system to manipulate the nutrient data in order to calculate nutrient values for intakes, menus, and recipes. These programs could be accessed either batch mode or on and on-line basis. We have continued to use these programs until the completion of our microcomputer program in 1985.

Maintaining the Data

Since 1973, OSUH Department of Nutrition and Dietetics has continued to maintain the NDB for internal use and for other users throughout the country. Dietitians have managed the project of data entry, selection and conversion of units. Yearly updates provided by USDA in either pamphlet, book, or most recently, magnetic tape or diskette format have been added to expand food items to approximately 9,200 items. Food manufacturers are continually contacted by mail or phone to try to keep up with new products. The popularity of fast foods has warranted the soliciting of nutritional data from these operations.

Missing Data

Nevertheless, we still receive criticism concerning lack of data in the areas of ethnic foods, enteral formulas, and vitamin and mineral supplements. We are trying to expand these data to the best of our ability.

The other criticism is with regards to "holes" in the data. These missing data are inevitable because:

a) The number of nutrient analysis studies reported are limited.
b) Most investigators do not analyze a food item for every nutrient in our data base. They usually analyze only the nutrients involved in their study.
c) Many manufacturers do not provide nutrient information for their products because of cost of analysis. We have seen an increase in the availability of manufacturers' data through the years due to food labeling and FDA requirements.

We do not impute values for missing data. However, with the number of food items available from which to choose, the researcher has the option of looking for a food item which has the most complete data for the particular nutrient under study.

Our data analysis reports do distinguish between true zero values and missing values.

The Microcomputer Nutrient Data Base System

When the decision was made to reprogram our Food Management system, we were looking at microcomputer hardware for data storage. However, with the advent of the fixed disk microcomputers, we saw the opportunity to gain an enormous amount of flexibility in terms of programming and system operations.

In terms of program design, we naturally wanted to integrate our NDB with our food system. We saw the need to automatically analyze patient recipes on a per serving basis when an ingredient or amount changed. We were able to incorporate yield factors involved in food preparation and cooking to allow for the accurate calculation of nutrient values. We also needed the ability to calculate both in-patient and out-patient food intakes as well as perform nutritional screenings and assessments.

We decided to delete some of the nutrients of the original 63 which were no longer of interest for nutrition researchers and add others. We now have 68 nutrient components in our system. Our nutrient data selection criteria essentially remains the same except that USDA Handbook revisions are now included.

The NDB microcomputer system was programmed using a Data Base Management System language called "The O'Hanlon Database Solution". We chose a Data Base Management System to insure flexibility in data handling as well as ease in programming. We no longer were limited in file space except by the size of our hard disk.

The programming language is very easy to use and therefor has saved us time in the development and debugging of the programs. One of the functions within the DBMS called "INQUIRE" allows us to compare data within a file so that we can produce "quick and dirty" reports within a matter of minutes. These report formats can be saved for future use. The DBMS allows us to access food items by name, partial name or code number. A next or previous

function allows us to scroll through the file if we are unable to find the item immediately.

System Reports

The NDB reports give us a wide range of information. One of our system reports prints a hard copy of the 68 nutrient values for any item on a per 100 gm basis. We can also get values based on the weight of the first household conversion code listed for each item.

To aid in data maintenance, we have the computer automatically convert nutrient data to 100 gram portions rather than our former method of hand calculation by conversion factor. This has greatly improved the accuracy of our data as well as the speed of data entry.

As previously mentioned, food items can be accessed and entered for nutrient analysis by item name, partial name, or code number. Amounts are entered by selecting any one of three household conversion codes. The gram weight of the item can also be used to input amounts.

The Analysis Report includes P:S and Ca:P ratios, percent of food and/or supplement from diet, percent protein, fat, and carbohydrate from diet, nutrient values per item per meal and total of those nutrients for a day. We are currently working on an enhancement to enable us to analyze mean values of diets for up to 99 days. The user can select 68 nutrients, any combination of nutrients or a single nutrient for analysis. The analysis data is compared to the RDA's and Maximum Dietary Levels.

The Percent Nutrient Report shows us the percent and number of records containing a specified nutrient. This lets us know how complete our data is for a specific nutrient and where we might want to concentrate our efforts for data searches.

The Specific Nutrient Catalog allows us to select up to 7 nutrients to produce a listing of data base items which contain that nutrient. These values are reported per 100 gram portions. This information has proven most helpful for dietitians or researchers interested in a particular nutrient.

The Nutritional Assessment Report includes fields to search for enteral formulas in the NDB. Protein, carbohydrate, fat, calories, and nitrogen ratios are calculated using the values from the NDB.

Applications

The primary applications of our data base programs at OSU Hospitals have been in the areas of patient care, therapeutic dietetics and nutritional research studies. Food intake analysis studies are used for our diabetic patients, on the Oncology unit, the Renal unit and for the Eating Disorders Clinic.

The clinical dietitians have access to information on the nutrient values of various recipes. This information is often used for patient education purposes during diet counseling. We also use the recipe analysis information for planning menus.

A subset of our NDB is used by the University in their CAI (Computer Assisted Instruction) programs for teaching medical and dietetic interns.

Conclusion

The OSU Hospitals has maintained and continually enhanced the Nutrient Data Base for 15 years. We believe that we have developed a state-of-the-art system to access this data and appreciate the opportunity to share this information with you at this conference.