NUTRIENT DATA BANKS - THE STATE OF THE ART

Introduction

Harold B. Houser

The state of the art of computerized nutrient data banks will be put into focus by presenting problems and issues related to one specific computerized nutrient data base. It may be presumptuous to confine observations of the state-of-the-art to only one of the 15 to 20 data bases developed in the United States, however, the current status of the HVH-CWRU Nutrient Data Base and its historical development probably reflect the problems and issues related to other computerized nutrient data bases. To insure a non-parochial view, the reactor panel this afternoon will discuss the state-of-the-art in the broader context of computerized nutrient data banks in general.

I will review briefly the history of the HVH-CWRU Nutrient Data Base, since I believe the current status is better understood if one knows how it was arrived at. Our first need for a nutrient data base arose in 1960 when we were studying persons with chronic illness. A large proportion of people admitted to Highland View Hospital had evidence for nutritional deficiency at the time of admission. In order to study chronically ill persons before they came into the hospital, we selected for study a group of people with multiple sclerosis who were living at home. We planned to determine their usual diet over a year and then translate what they ate into nutrients. It became apparent that if we could do the former, we could not do the latter. A suitable nutrient data base was not available in hard copy, let alone a computerized version. If we were to do the study we had to develop a nutrient data base and, because of the large numbers of individuals that we were studying over a year's time, we had to use a computer.

Starting with USDA Handbook 8, we searched the literature, checked with manufacturers and even set up our own kitchen laboratory in order to measure and weigh food; eventually we compiled a nutrient data base which met our needs. This early work and initial development of the data base was done by Ardyce Sorensen who, having worked with Bertha Burke, found that the analysis of 36 days of recorded intake from a lot of individuals was quite different from analyzing a long diet history from one individual.

Once we had suitable nutrient data, the next step was to organize it through a food identification code that would make the life of coders not too difficult and would also make some sense to computer systems and applications programmers and to the users of the analyzed data-nutritionists, statisticians and epidemiologists. Again we found we were on our own since we could not identify a ready source of information to help us. Next, under the direction of Arthur Littell, we developed computer programming that enabled us to enter food consumption data, access the computer stored nutrient data base, calculate and sum individual nutrients and percent RDAs, and print out the nutrient data in a useful format. While each person had a primary responsibility in the development, the total system was the result of input into all phases of development by a multidisciplinary team whose interests were the use of the data generated as well as generation and processing of data. This point is emphasized since we feel this approach is necessary to develop a flexible system responsive to the users.

The success of our initial efforts enabled us to do nutrient analyses on over 5,000 daily diet records from free-living persons without having to limit our analysis to short food lists or to substitutions or exchanges. A problem then, however, and still a problem but less so, was the limited number of nutrients about which we had information. Our first food table had information on only 8 nutrients in addition to calories, protein, fat, and carbohydrates. We did gain useful additional information by identifying protein, fat, and carbohydrates in all food items by animal or plant source. This proved useful in our studies of the relation of usual diet and fatty acid composition of subcutaneous tissue. Then, as now, information about the fatty acid composition of food was limited. Other early developers of food tables such as Margaret Moore, Mary Ellen Goodloe, and the Heart Disease Control Program, Bureau of State Services, USPHS, had the same problem. We solved the problem by assigning iodine numbers to our food groups and, owing to the computer, were able to assign an iodine index to a person's annual dietary intake.

By the middle 1960's we had developed a useful system which could be applied to the studies in diet and atherosclerotic heart disease, of high interest at that time. We tried to identify others with nutrient data bases in order to exchange information and to develop consensus on content and structure. Variability in results of studies should not be from differing methods of analyzing dietary data. In 1967 the Heart Disease Control Program, supporting many of the dietary studies, recognized this and asked us to develop a contract proposal for a National Food Table.

In a statement that now seems simplistic, we proposed to:

> "develop a food list that will identify and classify all known food items used by the United States population. A suitable code for the food list will be

developed along with an outline or model of the nutrient table. A comprehensive food list and code suitable for study of diet by cardiovascular disease researchers and other related research groups is not now available. Such a list and code with the necessary subsequent development of an accompanying nutrient table will provide a basis of comparability between studies that is presently impossible with the several different lists in nutrient tables now in use. A systematic numerical scheme will be devised to group foods into major categories and sub-groups assigned a code number for each food item to permit addition of new market foods. Sources of information will include existing food lists, food composition tables, research reports and commercial statements. A committee will be appointed to advise regarding the development of the food list and code and the size and structure of the nutrient table which would be used with the code."

The contract, approved and due to start on August 1, 1968, was cancelled on June 30, 1968. The Heart Disease Control Program now had a new director and he had decided that developing a nutrient data base or food table was of very low priority.

Where would we be today if the work had been carried out? Probably not here discussing the above issues. The cancellation presaged the next several years of difficulty in obtaining federal or other support for continued maintenance and development of our food table. The problem of financing was pointed out earlier today by Joan Karkeck and was a prominent part of the agenda for the first Nutrient Data Bank Conference.

During this period there was little exchange of information, or even awareness, among the several persons in the United States working at the common problem. This is illustrated by a letter I received from Arden Forrey in July, 1973. It turned out that we had been in competition for a sub-contract from the CAPO (Computer Aided Physician's Office Practice) project to develop a computerized method for diet instruction of patients. He inquired who were we?, what were we doing?, and suggested that it was time to develop a national working conference to provide exchange of information. Although the first such conference did not occur for another three years, I believe our exchange of correspondence and the identification of others with similar problems were the initiating factors in planning the first conference.

In the meantime, we needed to develop an extensive food table in relation to a proposal for the National Exercise and Heart Disease Study. Our original table was outdated. Somewhat pessimistically we again searched for a suitable table. Unable to identify any that suited our purpose we started over.

Although the dietary part of the National Exercise Study was

not funded, we had gotten well along in development of the new food table and decided to complete it; surely funding sources would see the need for working nutrient data bases. When the Diet and Cancer program came along it gave us renewed hope. However, we were not successful in obtaining federal funding for development of the food table itself. One review came back suggesting that the applicants were "just trying to play with a computer". As a result our data base was not developed with federal support, and I think this is probably true of most of the "private" nutrient data bases. They were developed to meet a specific need and their utility has become apparent to many other users. Interest and demand has heightened along with a certain reluctance to pay the developers for more than the reproduction cost.

In summary, the past nineteen years of our experience is probably mirrored by all developers of computerized nutrient data bases. First, the definition of a need; then the disappointing search for somebody else's answer to the need; the long and difficult development of the computerized system in a lonely environment; the erosion of interest and support for the development; the dogged persistence in the belief that sometime, somewhere the need for what you had developed would become apparent; and finally finding that there is interest in a forum to discuss the problems and issues relative to computerized nutrient data bases.