

Arnav Shah

ashah@ic.sunysb.edu
December 9, 2007

Michael Spittel, Ph.D.
Demographic and Behavioral Sciences Branch
National Institute of Child Health and Human Development
6100 Executive Boulevard, Room 8B07, MSC 7510
Bethesda, MD 20892-7510

Dear Mr. Spittel,

The following is in response to Grant Notice PAR-07-259 posted December 21st, 2006 with the National Institutes of Health about improving diet and physical activity assessment. This proposal focuses on the former issue: improving diet assessment; current methods in acquiring data are subject to a good deal of inaccuracy. In improving methods of collecting data on these matters, we would have more sound data available in analyzing the effects of certain diet habits on the lives of people. It is known that chronic diseases such as cancer, coronary heart disease, and obesity are affected by lifestyle factors such as diet. Having a richer source of information to work with has very extensive consequences: we can come up with ways for anyone to live healthier lives simply through changes in lifestyle. Included in this proposal is the summary below, some background information on current and past techniques used in assessing diets, my proposal of “product-style” categorization, the benefits of such a system, the procedure in implementing and making good use of this system, benefits and feasibility of such a system, objectives in implementing the system, and cost estimates.

Problem: Improving Diet Assessment

The grant request is to find new approaches or improvements on existing approaches in assessing statistics on diet of respondents in research. While current approaches have come a long way in acquiring information on diet within lifestyles, these methods are subject to significant inaccuracies. Reports from respondents are subject to many different sources of errors. Logging diet information tends to degrade over a longer period of time. Data from shorter periods of time may not accurately represent what happens over a longer period of time. In addition, respondents may change diets in order to make reporting easier. Also, respondents may underreport on their diet.

In measuring information on diet, I bring an approach that starts with a different view on classifying diet components. Instead of classifying foodstuffs by their label information, like “carbohydrates” or “vitamins”, foodstuffs should be classified by what they actually are commercially, for example, “Cheerios from General Mills”. A global database can be set up for all such food items, even using the existing UPC system, and additional data can be extracted from here. Responders would easily be able to report long term information by providing receipts, as they are already written records.

Background on the Assessment of Diet

Current approaches have come a long way in acquiring information on diet within lifestyles. There have been a few different methods but each has been subject to significant inaccuracies. Reports from respondents are subject to many different sources of errors. For example, one method is to have respondents log information on their diets. When this information is logged over a long period of time, there has been a noticeable degradation in the quality of the information in the logs over time. Also, responders may have changed their diet to better accommodate the logging process resulting in a less accurate picture. Another method of obtaining information is to conduct a 24-hour recall, where the respondents would report their diets from the past 24 hours. While this avoided the problem of log degradation, it was subject to other inaccuracies. For one thing, the data retrieved may not have accurately reflected a large time period, which is what is really desired. Also, there was a tendency of respondents not remembering information correctly, either leaving out part of their diet or not knowing portions. A problem that has affected many trials in the past has been the honesty among respondents. It appears that many respondents are often been guilty of underreporting their diet.

Proposal

My approach to solving the problem of accurately obtaining data begins with classifying diet components differently. A lot of the current methods try to classify things in categories such as “carbohydrates” and “vitamin A”, or “pasta” and “roast beef sandwich with American cheese”. The problem with the former is that these classifications may not be consistent. For example, the carbohydrates from one kind of food may be different than the ones from another, and there is no proper distinction made when data is recorded. The problem with the latter is that because there are known differences between different food of the same name (like pasta), the burden of the distinction is left on the respondent, resulting in inaccurate and inconsistent information. I propose that the components of a diet be classified by their commercial listing, for example, “Cheerios from General Mills”. This “product-style” classification should extend well to most food products, as even most meals are comprised from a combination of multiple products.

Benefits of the “Product-Style” Categorization

There are several advantages to this. On the side of the respondent, there is much more ease and accuracy in reporting information. Respondents will have access to some sort of receipt records and there should be very little components of diet that are erroneously not reported. Also, while it is true that using these types of records are infeasible for a short duration, they should work well in longer periods of time, which is more important in analysis anyway considering that health effects are the long term consequences of diet. On the side of the researchers and analyzers, a lot of the type of information from previous approaches can be extracted from this reporting. Most consumer food products have “Nutrition Facts” labels, and this information is readily and easily available from the product reporting of respondents.

Procedure for Obtaining and Storing Information

As with any project involving information, databases will be required. In the case of this project, several large databases will be required. We would need one database to contain all possible food products that may be reported. Alongside this, we would need nutrition information for every product. Another database would contain all reported respondent information. These databases should not be particularly difficult to build and connect because a lot of the infrastructure already exists in our consumer system. For example, we have the Universal Product Code (UPC) on most food products. This provides us the link between the consumer records of food purchases and the food products, with their nutritional information. Obtaining respondent information can be made easier by utilizing the existing “membership card” system that many supermarkets use. These systems track customer purchases and are an excellent source of information already in an organized manner. In cases where this system does not exist, such as restaurants and smaller grocery stores, we have receipts to fall back on. In almost every case, we will have a receipt record that contains all the food products, and is dated. In tagging nutritional information to products, most data should already exist. On packaged food products, we have the “Nutrition Facts” label which consists of statistics and an ingredients list. At restaurants, while nutritional information is not posted readily, it is still accessible through other means as they are required to make this information available.

Results and Feasibility of this Database System

The biggest benefit with this system is that it will greatly reduce the burden on the respondents, thus resulting in more accurate data. Because information can either be gathered automatically through purchase tracking, or recalled easily by the use of receipt records, there will be a much smaller chance of errors and misreporting. Better data analysis can occur with more sound data. With this database system, we have a powerful way to access information. For example, we can see diet patterns of a respondent over a long period of time, and see intake of various nutritional components. This can be compared with available information on respondent health. We can see what the healthiest respondents are eating and what the least-healthy respondents are eating. Being able to analyze data in this manner has many benefits, which extend to every one of us. We can gain a firm grasp on dietary habits that are truly nutritional. People can then make healthier choices in diet; as opposed to the confusing mess we currently have on determining what is healthy. Quality of life can be improved and chronic diseases can be prevented. But again, all this matters on the quality of the data, and the system I have proposed helps greatly to improve data quality.

Objectives for Getting this System Functional

Here is a list of the tasks that need to be done in order to get these databases set up:

- 1 – Begin the database with the UPC of all food products; this can be taken from already existing systems like those used in supermarkets.
- 2 – Establish nutritional information for all items in this database; this can be obtained from the food manufacturers or restaurants as they are required to make this information available.

- 3 – Gain access to existing systems of respondent information tracking; establish agreements with major supermarkets that use a “membership card” system.
- 4 – Link existing systems together, UPC and supermarket systems.
- 5 – Develop an input system to allow respondents to enter paper records, like receipts, into the database.

All of these tasks can occur concurrently. Progress on some tasks would allow for progress in other tasks.

Costs for Implementation

Monetary costs for this project should be on the low end because there is little need for physical resources. However, a good deal of time and effort will be required for the initial implementation. After this initial implementation is complete, much less work will be required to acquire information because the system will be very automatic. The burden on respondents will be lessened, as well as the work required by researchers.

Closing

Diet within one’s lifestyle has significant consequences over health. There is a lot we do not know about long term consequences of food that people consume. But the information we need is available. The system I have proposed can allow us to obtain this information easily and accurately. Researchers can analyze this information in a straightforward manner and be able to draw inferences on the effects of what people consume. Knowledge on what is healthy and what is unhealthy can be attained and people can choose to live better, healthier lives.

Bibliography:

Pollan, Michael. "Unhappy Meals." The New York Times. 28 January, 2007
<<http://www.nytimes.com/2007/01/28/magazine/28nutritionism.t.html> >

Taubes, Gary. "What if It's All Been a Big Fat Lie?" The New York Times. 7 July 2002
<<http://query.nytimes.com/gst/fullpage.html?sec=health&res=9F04E2D61F3EF934A35754C0A9>>