GUIDELINES FOR USE OF DIETARY INTAKE DATA

December 1986

Prepared for

CENTER FOR FOOD SAFETY AND APPLIED NUTRITION
FOOD AND DRUG ADMINISTRATION
DEPARTMENT OF HEALTH AND HUMAN SERVICES
WASHINGTON, D.C. 20204

under

Contract No. FDA 223-84-2059
GUIDELINES FOR USE OF DIETARY INTAKE DATA

December 1986

Prepared for
CENTER FOR FOOD SAFETY AND APPLIED NUTRITION
FOOD AND DRUG ADMINISTRATION
DEPARTMENT OF HEALTH AND HUMAN SERVICES
WASHINGTON, D.C. 20204

under
Contract No. FDA 223-84-2059

edited by
Sue Ann Anderson, Ph.D.

LIFE SCIENCES RESEARCH OFFICE
FEDERATION OF AMERICAN SOCIETIES
FOR EXPERIMENTAL BIOLOGY
9650 Rockville Pike
Bethesda, Maryland 20814
FOREWORD

The Life Sciences Research Office (LSRO), Federation of American Societies for Experimental Biology (FASEB), provides scientific assessments of topics in the biomedical sciences. Reports are based upon literature reviews and the scientific opinions of knowledgeable investigators engaged in work in specific areas of biology and medicine.

This report was developed for the Center for Food Safety and Applied Nutrition, Food and Drug Administration (FDA) in accordance with the provisions of Contract No. FDA 223-84-2059. It was edited by Sue Ann Anderson, Ph.D., Senior Staff Scientist, LSRO, FASEB, based on discussions of and materials drafted by the ad hoc Expert Panel on Guidelines for Use of Dietary Intake Data. Scientists selected as members of the Panel were chosen for their qualifications, experience, and judgment, with due consideration for balance and breadth in appropriate professional disciplines. Members of the Panel and others who assisted in the preparation of this report are listed in Chapter VIII.

In particular, the Panel and LSRO acknowledge the cooperation of scientific staff of the Center for Food Safety and Applied Nutrition, FDA, who provided information on the Agency's needs for dietary data.

The Expert Panel met four times between December 1985 and October 1986 to obtain background information, identify pertinent issues related to the interpretation of dietary data, and develop drafts of the report. Members of the Expert Panel reviewed each draft and the final report and provided additional documentation and viewpoints for incorporation into the final report. The Expert Panel and LSRO accept responsibility for the study conclusions and accuracy of the report; however, listing of these individuals in Chapter VIII does not imply that individual Panel members specifically endorse all statements in the report.

The final report was reviewed and approved by the LSRO Advisory Committee (which consists of representatives of each constituent Society of FASEB) under authority delegated by the Executive Committee of the Federation Board. Upon completion of these review procedures, the report was approved and transmitted to FDA by the Executive Director, FASEB.
While this is a report of the Federation of American Societies for Experimental Biology, it does not necessarily reflect the opinion of each individual member of the FASEB constituent Societies.

January 31, 1987

Kenneth D. Fisher, Ph.D.
Director
Life Sciences Research Office
EXECUTIVE SUMMARY

This report provides guidelines developed by an ad hoc Expert Panel for the interpretation of dietary intake data. It was prepared in response to a request from the Center for Food Safety and Applied Nutrition of the Food and Drug Administration (FDA) to provide a working reference for evaluating data on dietary intake. In the course of its deliberations, the Expert Panel identified five general areas that underlie the majority of studies that FDA is asked to evaluate.

1. Prevalence of consumption of particular levels of foods or food components;

2. Comparison of intakes of different groups within the U.S. population;

3. Time trends in consumption of foods or food components;

4. Relationship of intake of a food or food component to a given health outcome; and,

5. Prediction of change in a measurable outcome resulting from a change in dietary intake.

This report conveys the Expert Panel's perceptions of appropriate ways to interpret data relating to the first four of the five areas of study. A very extensive literature beyond the issues considered by the Expert Panel in preparing guidelines for the first four areas pertains to prediction of change in a measureable outcome resulting from a change in dietary intake. The members of the Expert Panel recognized that their suggestions regarding prediction of change would be inadequate and did not extend their deliberations into this area.

The ad hoc Expert Panel considered guidelines for two different situations: 1) design of de novo studies with defined purposes and 2) interpretation of secondary analyses made on existing databases. During the course of its discussions, the Panel recognized that guidelines cannot be formulated for all possible situations. Therefore, the guidelines are stated in general terms to allow latitude in their use. The Panel also recognized that interpretation of data may differ for scientific and policy-making purposes. For research purposes, it is often acceptable to conclude with statements that are "probably true" as a segment of a chain of scientific inquiry; however, for policy decisions, the best interpretation of available data must be used to make informed judgments for regulatory or legislative purposes. The text of this report
details the scientific rationale for interpretation of dietary data which should aid in bringing such data into clearer perspective for subsequent use as a part of policy decisions.

The Executive Summary presents the guidelines developed by the Expert Panel with references to appropriate discussions in the text of the report. General guidelines relevant to all of the areas of study identified above as concerns for FDA are listed in Section A, followed by guidelines specific for each particular area of study in Section B.

A. GENERAL GUIDELINES RELEVANT TO ALL AREAS OF STUDY

1. Definition of problem
   • The attribute or event of interest should be characterized as completely as possible. (p.3-20; 33-36)
   • The target population should be identified as definitively as possible. (p.23-28; 33-36)
   • The purposes for which the data are to be used and the precision needed should be determined. (p.3-15; 37-38; 48-49)
   • The adequacy of the dietary database for the food component(s) should be determined. (p.20-22; 57-60)

2. Sampling
   • The sample should represent the target population as closely as possible. A probability sample of the target population is best if the results are to be used for making inferences about the larger population. (p.23-26)
   • The critical elements of sampling are the design and the sample size. These elements will determine the precision of the estimates and the power of the statistical tests. (p.23-28; 37-48; 53-57)
   • Estimates based on small samples or weak experimental designs may have very wide confidence intervals which will probably limit their utility. Statistical tests
based on small samples or weak experimental design will have reduced power. (p.7-11; 13-14; 23-36; 37-48; 53-55; 61-63)

- Sample size is related to the variance associated with the estimate of the mean. Increasing sample size by increasing the number of subjects or the number of days of dietary data collection provides an approach for improving the precision of the estimate of the mean. (p.7-11; 13-14; 26-27; 37-48; 53-55)

- If a study design uses a probability sample, then the results are generalizable to the target population. If the sample is flawed or is not a probability sample, the results may still be generalizable but on the basis of expert judgment rather than on mathematical theory. In the latter situation, the presence of biases is a concern. (p.23-28)

- Independent corroboration of results is desirable if biases are or may be present. Independent corroboration of results is required if the generalizability of the results to the target population is based on expert judgment. (p.23-27)

3. Methods for collection of dietary intake data

- The data collection method should reliably measure the variable that the investigator is attempting to measure. If the method used results in collection of data that are imprecise, incomplete, insufficient, or inaccurate, then use of the data will be compromised and will probably lead to misinterpretation. (p.3-7; 29-30; 37-38; 48-63)

- With appropriate quality control, quantitative daily consumption methods (replicated for 2 or more days to provide estimates of variance ratios from within the study data), semiquantitative food frequency methods, or even specific questions may provide useful data. (p.3-11; 37-38; 53-57; 61-63)
Use of quantitative daily consumption methods is preferred when information is needed about actual intake. For estimates of usual intake, large intraindividual variability in day-to-day intake is the most serious concern with use of these methods, particularly when data are collected for only 1 day. Replicated collection of data with appropriate corrections may lower the total variance and lessen bias. Measures to control variability and bias are most effective if planned prior to data collection and should be sufficient to attain the necessary precision of the estimate. (p.8-9; 26-28; 37-42; 53-55)

Data collected by use of food frequency methods may provide a less accurate estimate of actual recent consumption than data collected by quantitative daily consumption methods, but intraindividual variation is a much less serious problem. Such data may be subject to a systematic bias of subject estimation of intake. The estimate may be either higher or lower than actual intake. (p.9-11; 37-38)

4. **Analysis of data**

Analysis of intake of a food component requires an adequate database for content of the substance in foods and a survey design and dietary method appropriate for estimating intake of foods containing the component. (p.20-22; 57-60)

The precision of the estimate is determined by the study design, the sample size, and the extent of variability in the data. (p.26-28; 37-48; 53-57)

Analysis should take into account the inherent biases and limitations of the study design and data collection methods. Adjustments and corrections should be appropriate for the database. (p.26-33; 37-63)

Even after adjustments and corrections, results must be assessed for adequacy of
individual assumptions and for any remaining biases before an analysis is accepted. (p.23-33; 37-63)

- In studies utilizing data collected for a single day, external estimates of variance ratios should be applied very cautiously, at least until more information is available on sources of intraindividual variation. (p.8-9; 26-27; 39-48; 53-55)

- Analysis of data collected in studies based on experimental designs employing complex sampling procedures requires the use of appropriate sample weighting factors. (p.23-27)

- Cautious interpretation is required for data based on a single-day's data collection or a food composition database judged to be inadequate. (p.8-9; 20-22; 29-30; 37-63)

- Interpretation of results should be based on biological and statistical significance rather than on statistical significance alone. (p.15-16; 65-70)

**B. GUIDELINES SPECIFIC FOR PARTICULAR AREAS OF STUDY**

1. **Prevalence of consumption of particular levels of foods or food components**
   - Use of either fixed cutoff points or a probability approach for making prevalence estimates should have a sound biological basis. (p.65-70)

   - In making prevalence estimates, generalization to the total population is often required; therefore, probability sampling is the method of choice. (p.23-28)

   - Estimation of centiles may be biased by the enlarged variance resulting from a large amount of intraindividual variation. Higher centiles (greater than the median) and lower centiles (less than the median) are overestimated and underestimated, respectively. (p.8-9; 39; 65-70)
A systematic bias in the data resulting in overestimation of intake shifts the entire distribution curve to the right. Prevalence will be overestimated from the upper tail of the curve and underestimated from the lower tail of the curve. Conversely, a systematic bias in the data resulting in underestimation of intake shifts the entire distribution curve to the left, producing an under- or overestimation of prevalence depending on which tail of the distribution is of interest. (p.39-42; 61-63; 65-70)

Per capita food availability data cannot be used for making prevalence estimates because distributions of intakes of individuals cannot be derived from these data. (p.16-20)

2. Comparison of intakes of different groups within the U.S. population

Probability sampling is desirable for estimating differences in absolute levels of intake for groups. However, when it can be assumed that the difference under investigation is independent of sample selection criteria, differences among groups may be representative of differences in the total population. Even when a sample is not representative of the total population (i.e., sampling only certain socioeconomic groups) differences among groups may still be important. (p.23-26)

Estimates of the means must be sufficiently precise to detect differences among the groups compared. Sample sizes must be large enough to ensure adequate precision in estimating means. (p.7-11; 23-28; 37-42)

If sample sizes are small, variances of sample means will be large. This may preclude detection of differences among means of groups. (p.7-11; 37-42)

Single-day dietary intake data may be adequate for comparisons of groups if the samples are sufficiently large. (p.7-9; 11-14; 37-42)
• Sampling and data collection methods should be equivalent for the groups being compared. (p.3-11; 23-28)

• If instruments are equivalent, food frequency data for dietary intake may be useful for comparison of group means because systematic biases will probably cancel out. However, instruments for all methods for collection of dietary intake data may be subject to different interpretation by different cultural groups. This type of bias will not cancel out in comparison of intakes among different cultural groups. Food frequency instruments are more subject to this problem than are quantitative daily consumption instruments. (p.9-11; 37-38; 53-63)

• Per capita food availability data cannot be used for comparison of intakes of groups because distributions of intakes of individuals within the groups cannot be derived from these data. (p.16-20)

3. Time trends in consumption of foods or food components

• The conceptual basis for variables should be constant over time. (p.49-53)

• Sampling procedures should be equivalent across all time points studied. (p.23-26; 49-53)

• Methods should be equivalent across all time points studied. (p.49-53)

• Values in food composition tables may change over time because analytical methods have changed or because the composition of certain foods actually has changed. Food composition databases used for data analysis should accurately represent the actual composition of foods available at each time point. (p.20-22; 49-53; 57-60)

• Changes over time may occur because the target population has changed over time. (p.49-53)
- Time trends observed over a short time interval should be interpreted cautiously until confirmed for a longer time span. Differences smaller than methodologic error may not be detectable, particularly if a time trend is based on only two time points. (p.49-53)

- Per capita food availability may be useful for estimating time trends in consumption of foods or food components. (p.16-20)

4. Relationship of intake of a food or food component to a given health outcome

- Independent, dependent, and associated or confounding variables must all be clearly defined. (p.31-32; 33-36; 65-66)

- Estimates of relationships tend to reflect the corresponding target population values even when probability sampling is not used. However, probability sampling provides a mathematical basis for generalization. (p.23-26; 33-36)

- The statistical considerations pertaining to comparison of groups and analysis of relationships also apply to intervention trials and clinical studies. (p.31-32; 33-36; 37-48; 61-63; 65-66)

a. Classification

- When dietary intake is used as the classification variable, the large amount of intraindividual variation in intake observed with use of quantitative daily consumption methods may result in a large likelihood of misclassification. Similarly, estimates made by use of frequency methods are sufficiently imprecise that classification by health outcome may result in fewer misclassifications. Therefore, use of intake of dietary components as independent variables should be done with caution. (p.42-47; 65-70)
b. **Bivariate analysis**

- In bivariate analyses (e.g., ANOVA, simple correlation, regression), misclassification of subjects attenuates associations. (p.42-47)

- Systematic bias has little effect on bivariate analyses. It is unlikely to attenuate the correlation coefficients in bivariate analyses but will yield an error in the intercepts in linear regressions. (p.42-47)

- In bivariate analyses, the impact of intrindividually variation and methodologic error varies among analyses. (p.42-47)

- Statistical correction of correlations and regressions is possible, but can be misleading. Therefore, it must be done with caution. (p.42-47)

c. **Multivariate analysis**

- Intraindividual variation results in serious distortion for multivariate analyses. Statistical correction is generally not feasible and an increased number of replicates is the only way to reduce the distortion. (p.47-48)

- Systematic bias has little effect on multivariate analyses. It will not preclude detection of relationships although the description of the relationship (e.g., intercept in multiple linear regression) will be in error. (p.47-48)

- Methodologic error in frequency methods can result in serious attenuation for multivariate analyses. Neither statistical corrections nor increasing the number of measurements reduces the error in the measures of association. (p.47-48)
d. **Factor analysis**

- Although factor analysis potentially provides a means to deal with multicollinearity in intake data, its value as an analytic tool for use with dietary intake data remains to be determined. (p.32-33)
# TABLE OF CONTENTS

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Foreword</td>
<td>iii</td>
</tr>
<tr>
<td>Executive Summary</td>
<td>v</td>
</tr>
<tr>
<td>I. Introduction</td>
<td>1</td>
</tr>
<tr>
<td>II. General Considerations for Dietary Data</td>
<td>3</td>
</tr>
<tr>
<td>A. Individual Intake Data</td>
<td>3</td>
</tr>
<tr>
<td>1. Methodological considerations</td>
<td>3</td>
</tr>
<tr>
<td>2. Estimation of specific parameters</td>
<td>7</td>
</tr>
<tr>
<td>3. Specification of dietary variables</td>
<td>11</td>
</tr>
<tr>
<td>4. Additional variables</td>
<td>14</td>
</tr>
<tr>
<td>5. Nutrient bioavailability</td>
<td>15</td>
</tr>
<tr>
<td>6. Dietary vs. nutritional status</td>
<td>15</td>
</tr>
<tr>
<td>B. Food Availability Data</td>
<td>16</td>
</tr>
<tr>
<td>1. National disappearance data</td>
<td>16</td>
</tr>
<tr>
<td>2. Total poundage of food additives</td>
<td>18</td>
</tr>
<tr>
<td>3. Marketplace disappearance data</td>
<td>18</td>
</tr>
<tr>
<td>4. Household food availability</td>
<td>19</td>
</tr>
<tr>
<td>C. Adequacy of Food Composition Databases</td>
<td>20</td>
</tr>
<tr>
<td>1. Databases for nutrients</td>
<td>20</td>
</tr>
<tr>
<td>2. Databases for nonnutrient components of foods</td>
<td>22</td>
</tr>
<tr>
<td>III. General Statistical Considerations</td>
<td>23</td>
</tr>
<tr>
<td>A. Sampling</td>
<td>23</td>
</tr>
<tr>
<td>B. Bias and Random Variation in Measurement</td>
<td>26</td>
</tr>
<tr>
<td>C. Computation of Variances in Complex Sampling Designs</td>
<td>27</td>
</tr>
<tr>
<td>D. Intraindividual Variation</td>
<td>28</td>
</tr>
</tbody>
</table>
E. Imputation of Missing Values .................. 29
F. Transformations and Unusual Distributions ... 31
G. The Effect of Multicollinearity on Statistical Estimation ................ 31
H. Factor Analysis Methods ...................... 32
I. Epidemiological Studies and Intervention Trials ........................................... 33
1. Epidemiological studies ..................... 33
2. Intervention trials ............................ 36

IV. Issues Related to Statistical Analysis of Dietary Intake Data ..................... 37
A. Effect of Methodology on Statistical Parameters ............................................ 37
B. Univariate Analysis ............................ 39
C. Bivariate Analysis .............................. 42
D. Multivariate Analysis .......................... 47
E. Pooling Data Across Studies .................. 48
F. Assessment of Dietary Time Trends ............ 49
1. General considerations ....................... 49
2. Statistical problems ......................... 51
G. Error Terms in Dietary Data and Their Impact in Analysis ....................... 53
1. Variance attributable to irregularity of consumption ............................... 55
2. Variance arising in the estimation of food intake ...................................... 56
3. Variance arising from food composition data and computation .................... 57
4. Implications for analysis of various types of variances ............................ 61
I. INTRODUCTION

Estimation of dietary intake and/or food available for consumption is a component of scientific inquiry in disciplines including nutrition, toxicology, and epidemiology. Such estimates, or values derived from the estimates, must often be considered together with other factors in making policy decisions regarding effects of diet or changes in diet on the health of the general population or subgroups of the population.

The Food and Drug Administration (FDA) is responsible for ensuring the safety and adequacy of the nation's food supply. To carry out this responsibility, FDA must monitor information from many sources and consider results from national surveys having dietary components, from studies providing dietary information on smaller or specific groups within the U.S. population, from clinical or experimental studies of physiologic effects of specific foods or food components, and from food availability data. Results of such studies have implications for FDA policy determinations in several areas such as cereal grain fortification, nutrition labeling, and food safety.

In order to have a working reference for evaluating data on dietary intake, the Center for Food Safety and Applied Nutrition of FDA requested that the Life Sciences Research Office (LSRO) of the Federation of American Societies of Experimental Biology (FASEB) undertake a study to develop guidelines for interpretation of dietary data. LSRO convened an ad hoc Expert Panel of scientists having expertise in disciplines related to issues pertinent to development of this type of reference. Members of this Panel are identified in Chapter VIII.

The Center for Food Safety and Applied Nutrition submitted its request for this study to LSRO as a Scope of Work consisting of the six questions listed in Appendix A. Subsequent to the development of this Scope of Work, the National Research Council (NRC) published its report, Nutrient Adequacy, (National Research Council, 1986) which dealt with parts of the questions specified in FDA's Scope of Work. In order to avoid duplication of effort, the ad hoc Expert Panel has referenced the NRC report and has interpreted the Scope of Work in terms of what it perceives as issues for FDA related to interpretation of dietary intake data.

In the course of its deliberations, the Expert Panel identified five general areas that underlie the majority of studies that FDA is asked to evaluate.
1. Prevalence of consumption of particular levels of foods or food components;
2. Comparison of intakes of different groups within the U.S. population;
3. Time trends in consumption of foods or food components;
4. Relationship of intake of a food or food component to a given health outcome; and,
5. Prediction of change in a measurable outcome resulting from a change in dietary intake.

To make these evaluations, FDA must often use previously collected databases to do secondary analyses for purposes that may be quite different from the purpose for which the data were originally collected. The Expert Panel has attempted to specify the type(s) of data needed for analyses related to such areas of study rather than to judge the appropriateness of use of particular databases for specific purposes of analysis.

The Panel recognized that data related or applied to diet are collected and analyzed for many purposes outside the scope of this study. In addressing the concerns of FDA, the Expert Panel has emphasized analysis and interpretation of the types of data that it considered more useful to FDA; however, this focus is not meant to imply lessened importance of dietary data collected for purposes outside the scope of this effort.

This report conveys the Expert Panel's perceptions of appropriate ways to interpret data relating to the first four of the five areas of study identified above. In this report the Panel has considered guidelines for two different situations: 1) for design of de novo studies for defined purposes, recognizing that the purpose and scope of the study will determine the most appropriate method for collection and analysis of data and 2) for interpretation of analyses made on existing databases, recognizing that the database may not have been designed for the analytic purposes at hand. The following sections of the report detail the methodological, statistical, and biological issues that the Panel considered important in analysis and interpretation of dietary data relating to the areas specified above.

The report first presents general information regarding dietary data collection and statistical concerns and then addresses specific issues important in interpretation of dietary data.
II. GENERAL CONSIDERATIONS FOR DIETARY DATA

Collection of information regarding diet has been approached from two directions. The first involves collection and analysis of dietary data from individuals. The second makes use of data on amounts of foods available to the total population or to particular units within the population, i.e., households. Use of these two approaches results in collection of diverse types of data that are appropriate for different purposes. Usefulness of data garnered from both approaches is considered by the Expert Panel in subsequent sections of this chapter.

A. INDIVIDUAL INTAKE DATA

For studies of dietary intake of individuals as well as for other investigations involving human subjects, the treatment of study participants, measurements, and interventions is of great importance for the subsequent interpretation of the data. The following questions indicate some general considerations for interpretation of empirical data. Are the correct items measured in the sense that the methods measure what is intended? Are all measurements made with adequate precision and with identical measurement procedures throughout the study? Are independent quality control procedures employed? Are data gatherers blinded in such a way that their biases and preconceptions cannot influence the outcome of their measurements? Data quality is best when all subjects are treated in exactly the same objective way.

1. Methodological considerations

For the purpose of this report, methodology for collecting dietary intake data from individuals is categorized as quantitative daily consumption methods and food frequency methods. These constructs provide a convenient means of classifying particular methods and instruments used for data collection; however, the Panel recognizes that the methods can be used to measure conceptually similar parameters depending, in large part, on the study design and protocol. Several investigators have reviewed the rationale, applications, and limitations of the various instruments commonly used to apply dietary intake methods (Burk and Pao, 1976; Keys, 1979; Marr, 1971; Pekkarinen, 1970). The methods and their respective instruments are operationally defined in the text as an orientation to the discussions in the remainder of the report. Characteristics of dietary intake methodology are summarized in Table 1. Selected biometric terms are defined in the Glossary of Statistical Terms (Appendix B).
<table>
<thead>
<tr>
<th></th>
<th>Quantitative Daily Consumption Methods</th>
<th>Food Frequency Methods</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Instrument</strong></td>
<td>Food record; food recall</td>
<td>Food frequency and diet history questionnaires</td>
</tr>
<tr>
<td><strong>Observation Unit</strong></td>
<td>Individual record of actual food consumption in a day or recall of food consumption during the previous 24 hours</td>
<td>Individual recall of usual dietary pattern during the target period of time</td>
</tr>
<tr>
<td><strong>Parameter Estimated</strong></td>
<td>Actual intake; may approximate usual intake if number of replicates is sufficient</td>
<td>Usual intake</td>
</tr>
<tr>
<td><strong>Estimation Method</strong></td>
<td>True mean intake of a dietary factor for an individual is estimated from a number of individual measurements</td>
<td>True mean intake by an individual is estimated directly</td>
</tr>
<tr>
<td><strong>Usual Timespan Targeted</strong></td>
<td>Shorter interval</td>
<td>Longer interval</td>
</tr>
<tr>
<td><strong>Relative Accuracy of Data for Specific Days</strong></td>
<td>More accurate</td>
<td>Less accurate</td>
</tr>
<tr>
<td><strong>Meaningful Absolute Value</strong></td>
<td>Yes</td>
<td>Less absolute</td>
</tr>
<tr>
<td><strong>Preserving Rank Order</strong></td>
<td>Yes, if a large number of replicates is used</td>
<td>Maybe</td>
</tr>
<tr>
<td><strong>Problems Caused by Intraindividual Variation</strong></td>
<td>Yes</td>
<td>Less serious</td>
</tr>
<tr>
<td><strong>Seasonal Variation Taken into Account</strong></td>
<td>Yes, if replicates are collected in different seasons</td>
<td>Yes, if designed properly</td>
</tr>
<tr>
<td><strong>Inclusion of Infrequently Consumed Foods</strong></td>
<td>Likely to miss if the number of replicates is small</td>
<td>Likely to include if list includes particular food(s)</td>
</tr>
<tr>
<td><strong>Relative Feasibility</strong></td>
<td>More difficult if a large number of measurements is needed.</td>
<td>Easier</td>
</tr>
<tr>
<td><strong>Major Problems</strong></td>
<td>Intraindividual variability and inadequate number of measurements</td>
<td>Inaccuracy of estimates of intake</td>
</tr>
</tbody>
</table>
a. Quantitative daily consumption methods

The distinguishing feature of quantitative daily consumption methods is that they attempt to capture the exact nature and quantity of individual foods consumed in defined periods of time. This contrasts with the food frequency methods which attempt to capture information about patterns of food use and implied nutrient intake across longer, but necessarily less precisely defined, periods of time.

The term "quantitative daily consumption method", as used in this report, refers to the basic unit of time (1 day) for which data are collected when this method is used. The Panel is concerned that use of the terms "recalls" and "records" often implies different periods of data collection. Reports in the literature often refer to "24-hour recalls" in the connotation of collection of information about a single day's intake and "food records" with the implication of several days' intake. However, food records may be kept for only 1 day and food recalls may be repeated to supply many days' data. The Panel considers the number of days that data are collected for each subject to be a more important focus than the specific instrument employed. Therefore, the discussions of this report focus on the interpretation of quantitative daily consumption data in terms of total time periods for which data are collected rather than the particular instrument employed.

Instruments most frequently used to collect dietary data by quantitative daily consumption methods are recalls and records of actual food intake over a finite period of time. For food recalls, subjects list from memory food intake within a specified period, usually the immediate past 24 hours, and estimate quantities of each food consumed, usually in household measures or servings of specified size. Either subjects or trained interviewers may record the data. For food records, subjects or observers are instructed to record foods and quantities eaten at the time of consumption. Quantities consumed may be estimated or measured in household units or by weight. Probes and food models may be used with either instrument to obtain more complete information and more accurate estimates of quantities eaten (Burk and Pao, 1976; Fehily, 1983). Particular protocols used with these instruments may have significant effects on the accuracy of the data collected and on the analysis and interpretation of dietary intake data as discussed in Chapter IV.

b. Food frequency methods

The term "food frequency" encompasses a family of methods that includes all instruments for which subjects recall their usual dietary intakes during a targeted time period (of more than 1 day) in the past (Liu, 1986).
These methods are most often used to estimate usual intake of foods or classes of foods and of food components in order to rank or categorize subjects rather than supply quantitative measures of actual intake (Block, 1982). Data collection instruments include general and focused food frequency questionnaires, and diet histories.

Food frequency questionnaires consist of a self- or interviewer-administered list of foods and frequency-of-use response categories. Lists of foods for food frequency questionnaires may be extensive in order to provide estimates of total dietary intake or may be focused on groups of foods or particular foods that are good sources of specific components. The latter type of questionnaire is sometimes referred to as "focused questions"; an example is given by Heller et al. (1981). Additionally, specific questions may be directed toward particular aspects of food intake, such as use of alcohol or condiments or eating or avoiding visible fat, which might otherwise be missed.

Terminology regarding quantitation of food frequency questionnaires differs among investigators. The terminology of Sampson (1985) is used in this report. Questionnaires are considered qualitative if information on portion sizes is not provided and semiquantitative if portion sizes are specified. Levels of quantitation differ among questionnaires and may affect estimates of intake. Issues related to quantitation of intake from food frequency data are reviewed by Sampson (1985).

The diet history instrument also describes usual food intake over a relatively long time period. The research dietary history was developed originally by Burke (1947) to relate long-time customary food habits to clinical measurements. As administered by Burke (1947), this instrument consisted of two parts: 1) a 24-hour recall of actual intake and collection of information to estimate usual intake; and 2) a cross-check to verify information provided in the first part. The research diet history, as developed by Burke (1947), has not been used in large epidemiologic studies or surveys because of the time and cost associated with its use. In current usage the terms "diet history" and "food frequency" have become virtually synonymous and are often used interchangeably.

The distinctions between quantitative daily consumption methods and food frequency methods have important implications for data analysis and interpretation. In essence, there is a trade-off between error terms in the choice of methods. Given that the choice of method also carries important logistic and cost implications, these factors must be considered concurrently with the effects of the error terms in design of studies.
c. Appropriate use of methods

Appropriate interpretation of information on dietary intake depends on the reliability and validity of the methods and the degree to which they are appropriate for the purpose of the study. The report of Burk and Pao (1976) includes a comprehensive discussion of aspects of reliability and validity of dietary methods.

Reliability of dietary intake instruments has often been estimated by comparing results generated at two different times by the same instrument. Reliability assessment should be done cautiously since no exact measurement of reproducibility is possible. Agreement between two records may indicate an unvarying diet and an accurate instrument or it may indicate the tendency of subjects to give the same responses over time. Disagreement between two records may indicate an actual difference in diet and a reliable instrument or, conversely, a consistent diet and an unreliable instrument (Block, 1982).

Two concepts are important in the establishment of the validity of a dietary method and its instruments. First, what does the method purport to measure? Except for direct observation, there is no absolute measure of consumption that can be used as a standard of comparison for data obtained by the various dietary intake instruments (Burk and Pao, 1976). A few studies such as those of Madden et al. (1976) and Gersovitz et al. (1978) have used unobtrusive observation techniques to validate methods, but most studies of the validity of dietary intake methods have used one instrument of unknown validity as the standard for evaluation of the validity of a second instrument. Second, does the method suit the purpose of the study? As will be discussed extensively in the report, most relationships between diet and a biological outcome depend on intakes persisting over moderate periods of time ("usual intakes"). A reliable measure of short-term intake may have little validity or reliability as a descriptor of "usual intake" (Beaton, 1982).

2. Estimation of specific parameters from dietary intake data

The term "food intake" is used in two contexts in the literature on food consumption. Food intake may refer to estimates of actual recent intake during a finite period of time, for example, estimates of 1, 3, or 7 days. The same term may also refer to estimates of usual intake over a period of several weeks or months. Data on actual intake consist of detailed information on specific foods eaten, whereas data on usual intake comprise more general information regarding foods consumed. The two parameters are very different and data
regarding one parameter should not be assumed to have a particular relation to the other parameter. For example, actual intake measured very accurately on a single day or even on multiple days has been considered to provide a poor estimate of usual intake. However, there are studies in progress to investigate correction of actual intake data as a means of estimating usual intake (Rosner and Willett, 1987).

a. Quantitative daily consumption methods

The quantitative daily consumption methods are advantageous in that they can provide accurate information on the actual intakes of foods or food components. Such information may be needed for determination of an immediate physiologic or toxicologic effect. Distributions of such data are useful for estimation of the mean, median, and quantiles of intakes for groups of individuals who share some qualities such as age, sex, race, education, or income. Additional information that may be useful in interpretation of dietary data is listed in Section A-4 of this chapter.

The major difficulty encountered in analysis of dietary data obtained by quantitative daily intake methods is the large intraindividual variation inherently present. Unless diets are monotonous, day-to-day food intake differs more within each person (intraindividual variation) than it varies between persons (interindividual variation). Ratios of intraindividual variation based on data collected by quantitative daily consumption methods for several studies may differ among food components; among age, sex, and sociocultural groups; and within and between dietary intake instruments (recalls and records) (National Research Council, 1986). (See Chapters III and IV for discussion of effects of intraindividual variability on study design and analysis and interpretation of dietary data.)

The number of days for which records or recalls are obtained influences the appropriate use of the data. A single day's data are not adequate for evaluating dietary intake of individuals (Beaton et al., 1983; Liu et al., 1978; Todd et al., 1983; Young, 1981). Data obtained by dietary recall or record for a single day may reasonably estimate the mean intake for a population if the sample size is sufficiently large to minimize chance errors (Beaton et al., 1983; Liu et al., 1978; Todd et al., 1983); however, data based on a single day's data typically result in a distribution that is flatter and wider than the population it represents. Prevalence of high or low intakes is overestimated when estimates are based on such a distribution (Hegsted, 1972). Similarly, the power to detect differences in mean intakes of groups is lower when the data contain a large amount of intraindividual variation. As discussed in Chapter IV, large intraindividual variation associated with single-day
dietary data may mask associations between dietary intake and disease (Beaton et al., 1979; Jacobs et al., 1979; Keys et al., 1966; Liu et al., 1978); consequently, the Panel regards studies in which only 1 day's dietary information is obtained to be less preferable, in most cases, for examination of relationships between diet and disease.

Collection of data for multiple days provides a means of decreasing intraindividual variability, thereby decreasing the width of the confidence interval around the mean and increasing the precision of intake estimates. As illustrated by Beaton et al. (1979), the width of the confidence interval for the group mean may be narrowed by increasing the sample size or by increasing the number of days of replicate observations. Conversely, when it is necessary to estimate the average intake of each individual over time, the only approach to improving the reliability (reducing the width of the confidence interval) is to increase the number of days of data collection (Beaton et al., 1979).

When data are collected for multiple days, study design can minimize effects of day(s) of the week, season(s) of the year, and correlations between adjacent days of data collection (National Research Council, 1986). However, each added day of data collection provides a progressively smaller increment of new, independent information. Intraindividual variation differs greatly among food components; and thus, the rate of decrease in intraindividual variation resulting from additional days of data collection also varies substantially (Morgan et al., 1987a; Sempos et al., 1985). The optimal number of days for collecting intake data depends upon the purpose of the study, the precision of the estimate required, the food component(s) of interest, and the amount of intra- and interindividual variation (Morgan et al., 1987a; Woteki, 1986). For studies relating dietary intake to biological outcomes, it is critical that the dietary data and the biological data relate to the same timeframes (Underwood, 1986). Thus, regardless of methodologic choice, it is important to ask what period of time is represented as well as how reliably and validly the intake during this period is represented.

b. Food frequency methods

Data collected by food frequency methods are considered more representative of subjects' usual diets than actual intake data collected for a single day or only a few days; however, accuracy of the estimate is sacrificed. This loss of accuracy may or may not be critical, depending on the purpose of the study. Foods habitually eaten, foods eaten very infrequently, or those eaten in association with special events appear to be estimated most successfully by these methods. Data obtained by semiquantitative food frequency instruments
may be used to identify extremely high or low levels of intake by ranking the intakes of specific foods or food components (Samet et al., 1984). In interpreting data collected by semi-quantitative food frequency instruments, it should be recognized that collection of data by these instruments may result in a systematic bias toward either under- or overestimation of quantities, in contrast to the tendency for underreporting with the food record or recall instruments (Marr, 1971). Use of food frequency methods is being investigated as a means of estimating food intakes at times in the remote past; however, results obtained to date are inconclusive (Byers et al., 1983; McDonald et al., 1986; Van Staveren et al., 1986).

Questionnaires focused on specific items may provide information on usual intake of specific foods or food components. In some cases, relatively short lists of food items may well represent much of the American dietary over an extended time period (Block et al., 1986; Byers et al., 1985; Johnson et al., 1974; Murphy and Calloway, 1986; Willett et al., 1985).

Accuracy of estimation of food intake is affected by the manner in which amounts are quantified when food frequency methods are used. Food intake data may be aggregated in one or more stages. This process may be complex and the accuracy of each stage and the total process is not known. Significant error may be associated with respondent aggregation of intake over periods of high and low intake.

Use of food composition data in analysis of studies utilizing food frequency methods may involve grouping together of foods having similar contents of certain nutrients or other food components. Composition values for groups of foods in current databases are weighted averages of the contents of the individual foods included in the group. These weighted averages can be used to estimate the intake of a food component if the content of that component in each food in the group is randomly distributed around the mean; however, if this is not the case, calculation of intake from such weighted averages may result in significant bias (National Research Council, 1986). Grouping of data additionally assumes consumption of certain portions and proportions of items within each group. If this does not represent the true situation, use of weighted averages may introduce significant bias (Guthrie et al., 1984).

From the above, it is clear that use of food frequency methods carries an implicit loss of precision of estimation of actual consumption of foods and food components in comparison to the quantitative daily consumption methods. Conversely, food frequency methods are much less affected by intraindividual variation since they explicitly attempt to estimate the average intake across longer periods of time, often weeks or months (Beaton, 1982; Willett et al., 1985).
As the number of days of information included in quantitative daily consumption methods increases or as the specification of portion size and food items increases in food frequency methods, the results come closer.

3. Specification of dietary variables

Specification of variables is the step in the analytical process that precisely defines the elements to be studied. Variables can be broadly categorized as intakes of food components, specific foods, and groups of foods sharing a common characteristic. Examples of food components include "generally recognized as safe" (GRAS) food ingredients, food additives, nutrients or other substances occurring naturally in foods, and contaminants. Groups of foods having a common characteristic might include breakfast cereals or fruit juices. Examples of use of foods or food groups by populations or population subgroups include the proportion of persons consuming green leafy vegetables daily or the proportion of children consuming ready-to-eat cereals.

a. Food components

Analysis of dietary intake data to estimate consumption of food components requires consideration of the distribution of the component in foods, the pattern of consumption, and the proportion of the population or subgroup(s) consuming foods containing the component of interest. The Panel discussed three subcategories of consumption of food components based on these factors. The first subcategory consists of components that are widely distributed in foods and regularly consumed by all persons in the population. Examples include protein, fat, carbohydrate, and total energy intakes. The second subcategory is made up of components found in a limited number of foods, not consumed by all persons, and for which consumption may be periodic. Examples include artificial sweeteners and caffeine. The third subcategory consists of components concentrated in certain foods or food groups, consumed by essentially all persons, but for which consumption is sporadic. Examples include vitamin A, calcium, or a pesticide used only on fruit. Considerations for use of quantitative daily consumption and food frequency methods are detailed in the paragraphs that follow.

(i) Food components widely distributed in foods and regularly consumed by all persons. A single day's recall or record may accurately estimate the group mean intake if the sample size is sufficiently large. However, the amount
of intraindividual variability inherent in a single day's intake is large. Collection of dietary data for multiple days will lessen the intraindividual variability and result in a more precise estimate of intake.

Intakes of food components widely distributed in foods and regularly consumed by all persons may be estimated from frequency data, but the estimate may be less precise than an estimate based on quantitative daily consumption data.

(ii) Food components found in a limited number of foods, not consumed by all persons, and for which consumption may be periodic. Single-day measures of food intake may miss the days that foods containing these substances are consumed. Because the component is contained only in certain foods and because only a portion of individuals in the population will consume those foods on any given day, the number of respondents naming those foods is most likely to be substantially smaller than the total sample. The same considerations cited in the preceding section hold for the use of single and multiple days' data; however, for these substances, larger samples or disproportionate sampling of particular subgroups and collection of data for a greater number of days may be required. Although subgroups most likely to consume the food component of interest should be included in the sample, it should be recognized that over- or undersampling of these subgroups may bias results unless analyses are properly weighted. Analyses of intakes of "eaters only" may be preferred for such substances.

Intakes of components found in a limited number of foods, not consumed by all persons, and for which consumption may be periodic, may be better estimated by use of semiquantitative food frequency instruments than by quantitative daily consumption instruments. Focusing a food frequency questionnaire on foods that contain the component of interest may be the most efficient approach for estimating intake.

(iii) Food components concentrated in certain foods or food groups and consumed by essentially all persons, but for which consumption is sporadic. Single-day data will again likely miss the days that individuals consume such substances. In these cases, use of multiple days' data, distributed to include all days of the week and the specific parts of the year that such components are likely to be consumed, would increase the likelihood that a larger and more representative proportion of the sample would consume the component of interest. The foregoing discussions refer to the reliability of either the group mean intake or the mean intake of each individual across time. When analytical interest focuses upon the distribution of usual intakes among individuals, without the need to know
each individual's usual intake, a procedure has been described for the adjustment of observed distributions, using information derived from a minimal number of replicated 1-day intake estimates (National Research Council, 1986).

Intakes of components concentrated in certain foods or food groups and consumed by essentially all persons, but for which consumption is sporadic, may be better estimated by semiquantitative frequency instruments than by quantitative daily consumption instruments. Focusing a food frequency questionnaire on foods that contain the component of interest may be the most efficient method of estimating intake.

(iv) Nonfood sources of food components.
While the need collect data on food sources of food components is obvious, exposure to nondietary sources of substances such as drinking water, nutrient supplements, medications including antacids, colors used in cosmetics as well as foods, and contaminants including pesticide residues or industrial chemicals may also need to be taken into account. For example, the study of Sowers and Wallace (1986) illustrates the contribution of drinking water to total calcium intake in geographic areas having hard water. Because data from several surveys indicate that significant proportions of the U.S. population use dietary supplements (McDonald et al., 1986), the Expert Panel considers that dietary data gathered for the purpose of estimating nutrient intakes ideally should include complete information on brand, composition, and amounts of supplements used. If this information is omitted, a bias towards underestimation of nutrient intake will result. If the information is incompletely recorded, an error of unknown nature and magnitude will result. When data are collected on supplement use, intake data may be analyzed with or without the contributions made by the supplements depending on the purpose of the analysis.

b. Foods and food groups

(i) Quantitative daily consumption methods.
Estimation of intake of foods or groups of foods is equally or more difficult than estimation of intake of food components. When information is needed on consumption of particular food items or groups of foods, suitability of dietary intake data collected by quantitative daily consumption methods depends in large part on the pattern of consumption, the proportion of the population consuming the foods, and on the time period covered by the data collection.

Classes of foods (i.e., various green leafy vegetables) tend to be substituted for one another from day to day, making intraindividual variation greater for intake of specific foods
and food groups than for food components. Sempos et al. (1985, 1986) have found more variability in intake of some food groups than nutrients. Additionally, intraindividual variation is much larger than interindividual variation in intake of specific foods (Beaton, 1982; Sempos et al., 1985, 1986). Therefore, single-day data are not very useful for estimating an individual's usual intake of specific foods or food groups unless the pattern of food use is very consistent within individuals. Single-day data may suffice for estimates of group intakes of foods or food groups if sample size is sufficiently large or if food groupings are broad. As in the case of food components, the precision of the estimate will be improved by increasing the sample size or the number of days of data collection. It follows also that variability of intake of single foods or food groups in a population will be overestimated by single-day data.

(ii) Food frequency methods. Intake of specific foods or food groups is probably best estimated by food frequency methods because use of these methods permits representation of food intake over longer periods of time than quantitative daily measures of consumption for single or multiple days. Focused questions may be most efficient for this purpose, particularly when the food is consumed periodically, i.e., alcohol on weekends or fruit only in season.

4. Additional variables

Analysis of dietary data must include additional information collected with the dietary data set. Interpretation subsequently depends on an assessment of the quality and meaning of these other variables. Some variables of interest to FDA are listed below.

In population studies

- Demographic descriptors including age, race, sex, height, weight, pregnancy or lactation, etc.

- Geographic distribution, particularly if there is suggestion of geographic differences in trace element composition or contaminant distribution, etc. This assumes that there are region-specific data on food composition.

- Socioeconomic variables, including a review of the scaling system to ensure that it is scaling the same dimensions as other conventional population data.

- Seasonal variation in food intake.
In clinical studies

- Most variables listed above for population studies, in addition to the following.
- Criteria for inclusion and exclusion; for example, diagnostic criteria for specified disease conditions.
- Descriptors of current clinical state.
- Compliance measures in intervention studies.
- Response rate and treatment of data for nonrespondents.

It is not possible to establish objective criteria for the assessment of this information. Its adequacy depends upon the question being asked and the inference to be made and requires judgment by the analyst and interpreter. All that can be indicated here is a reminder that data of the type mentioned should be available and should be assessed.

5. Nutrient bioavailability

Interpretation of dietary intake data is further complicated because differences in composition of food mixtures consumed affect the amount of nutrients or other food components available for absorption. For most nutrients, a number of factors affect bioavailability. These include the forms of the nutrient in foods (e.g., folate monoglutamates vs. polyglutamates), other substances in foods that affect absorption (e.g., vitamin C enhances and phytates reduce iron absorption), and other substances required for absorption of a nutrient (e.g., necessity of zinc for activity of folate conjugase in the small intestine). Present food composition tables do not provide all of the information that would be required to calculate nutrient bioavailability on a meal-by-meal basis; further, even if the compositional data were available, there is not yet agreement on effective algorithms to make such calculations (National Research Council, 1986). It should be recognized that this source of variation is uncontrolled in most data analyses and that its impact on interpretation of such analyses will depend on the nutrient in question and on the exact purpose and nature of the analysis.

6. Dietary vs. nutritional status

An assessment of dietary status of an individual or a group of people does not immediately translate to an assessment of nutritional status (Beaton, 1982). There is a series of processes that occur between consumption of a nutrient
or food component and its final impact on the human body. Absorption, transport, storage, mobilization, and metabolism are physiological processes and nutrient status is the result of these intertwining and related sequences of events. In order to determine nutritional status, measures of biochemical indices, clinical observations, and anthropometric measurements are needed. This report considers dietary status which can lead to maintenance or changes in nutritional status; however, it is not the sole factor affecting the nutritional condition of an individual.

B. FOOD AVAILABILITY DATA

For the purpose of this report, the term "food availability data" refers to food/dietary component/food additive survey data that have not utilized the individual as the unit of observation. Rather, the surveyed populations are households or national or regional food supplies. Techniques for data collection and interpretation of results from these types of surveys are highly specialized and have not been considered in depth by the Expert Panel.

1. National disappearance data

The Economics, Statistics, and Cooperative Services (ESCS), U.S. Department of Agriculture (USDA), compiles annual supply and use data for most major agricultural commodities disappearing into consumption. The annual supply of each food consists of stocks carried over from the previous year, total production in the United States, and all imports and shipments from territories. Utilization consists of exports, shipments to territories, government purchases for military use and exports, nonfood use, and food use. Civilian food use is derived as a residual after deducting the other uses and the ending stocks from the total supply and, thus, is subject to the net errors in the other components. Civilian per capita disappearance is subsequently calculated by dividing total disappearance by the number of people eating from the civilian supplies.

The supply and utilization data system for food products depends entirely on data collected for other purposes. In addition, use of different foods is estimated at different stages in the food distribution system because of the complexity of the system. Consequently, food availability data contain significant gaps that become progressively more serious approaching the level of the consumer. For example, manufacture of primary products is covered relatively well, but manufacture of secondary products is covered much less
completely and retail manufacture is not reported at all. These constraints and others were detailed by Manchester and Farrell (1981).

The nutritive value of the U.S. per capita food supply is estimated from the amounts of food energy and 15 nutrients available from 350 foods in the supplies moving to civilian consumers. Energy and nutrient contents are calculated from USDA food composition data (Burk and Pao, 1976). Quantities of foods reported as available for consumption represent foods "as purchased" and include refuse such as bones, rinds, and peelings. Estimates of the nutrient content of the food supply include nutrients from only the edible parts of foods, including parts of certain foods that are edible but not always eaten. For example, the nutrient values for meat include all of the separable fat that is usually left on retail cuts. Nutrient estimates also include food waste and nutrient losses that may occur in processing, marketing, cooking, and from plate waste. Some sources of nutrients, such as alcoholic beverages, are excluded from estimates. Though vitamin/mineral supplements in tablet form are excluded, vitamins and minerals added to enrich or fortify foods are included.

National disappearance data provide the only measure of trends in the levels of nutrients in the U.S. food supply over the past 75 years. The mean levels most likely overestimate actual consumption levels but such overestimates are assumed to be consistent across time because, in most cases, methodology has remained consistent (Welsh and Marston, 1982). However, some changes in some methods over time are recognized, for example, in estimation of disappearance of food fats and oils (Call and Sánchez, 1967). Since it is expected that different categories of individuals (e.g., children vs. adults) will consume different amounts of food measured, for example, as total energy intake, and since per capita estimates are calculated per person without adjustment, it must be expected that comparisons of per capita data across time or between populations will reflect any differences in demographic structure. This, too, must be taken into account in interpretation.

Disappearance data are used by USDA to assess the overall sufficiency of the foods available to the U.S. civilian population. This time series does not account for differences in the distribution of food among individuals or for losses in quantities of foods or nutrients, but these data do provide an estimate of the quality of the U.S. diet regarding specific nutrients. Although other nutrition-monitoring methods measure nutrient ingestion by individuals, the historical nutrient series is a rapid and inexpensive indicator of the general diet quality. These data provide a mechanism for evaluating the relationship between nutrient availability and agricultural
commodities and can be used to evaluate and predict effects of the technological, demographic, political, and socioeconomic changes in the U.S. diet. Considerable care must be exercised in any attempt to link trends in disappearance data to trends in disease or mortality (see discussion, Chapter IV).

2. Total poundage of food additives

The Committee on Food Additives Survey Data of the National Academy of Sciences/National Research Council (NAS/NRC) has collected food additives poundage data in 1970, 1975, 1977, and 1982 (National Research Council, 1972, 1979). These survey data represent estimated total poundage of each additive used in food annually in the U.S. food supply. The NAS/NRC Committee, under contract with FDA, collects these data from the food industry by mailed questionnaire. The uncertainty inherent in such data, because of both the voluntary nature of industry data submissions to NAS and the great complexity of the U.S. food supply, is recognized by NAS and FDA. Nevertheless, these data have been used, with caveats, to estimate human exposure to the food additives in the database.

The value assigned to dietary exposure to a food additive from total poundage data amounts to a calculation of per capita exposure based on an assumed total population of consumers. Analysis of the entire body of NAS data shows that the total poundage figures generally may tend to underestimate actual exposure (Rulis et al., 1984).

3. Marketplace disappearance data

Disappearance of food from the marketplace is measured by several commercial surveys. Market researchers have used both record and recall procedures to obtain information on food purchases, particularly for a limited number of food items. Their emphasis has often been on quantities purchased or expenditures for particular commodities, including information on brands. The primary purpose of these surveys is to meet the information needs of food processing and marketing companies and their advertising agencies (Abrams, 1981), although they have been used by agencies such as FDA.

Commercial surveys generally collect data from consumers who return questionnaires by mail. Most market research panelists are members of middle or upper-middle class households. Members of illiterate households and persons living in an institutional setting are excluded from the sample frame (Abrams, 1981). Thus, these data contain inherent bias and probably are not truly representative of the U.S. population.
The databases are structured to meet the requirements of specific marketing studies and to maintain confidentiality of data regarding use of specific products. Therefore, restrictions placed on the use of these types of market data may prevent public use of the data. Market research data provide a source of data for determining the percentage of individuals consuming particular food item(s). The broader interpretation of analysis of such data is limited by the proprietary nature of the data and the sample bias.

Records of disposition of food products within households are also collected periodically by some commercial surveys. Such records indicate the foods consumed, the consumer(s), and in some cases, information on quantities consumed. Foods consumed away from home by members of the household other than the person completing the questionnaire may be over- or underestimated. These "use data" have been used in conjunction with an estimated portion size of a food assumed to be consumed during a single eating occasion. Although the data are not fully disaggregated, such data sources, used in conjunction with a nutrient database and/or a database for food additives, have been used to estimate mean and percentile levels of intake of particular dietary components or food additives for different age groups (Abrams, 1981; Rulis et al., 1984). Assumption that average servings are consumed may introduce considerable error in the estimation of "intake" from food availability data.

4. Household food availability

The U.S. Department of Agriculture has conducted cross-sectional surveys of household food use in 1936, 1942, 1948, 1955-56, and 1977-78. Sampling procedures for household data collection are such that results are representative of U.S. populations when appropriate weights are used with these data. These surveys provide detailed information on food used by households from which the nutrient content of household food supplies can be estimated (Burk and Pao, 1976). Discrepancies should be expected between disappearance data calculated from national and household sources because food use is estimated at widely differing points in the distribution system (Food and Agricultural Organization of the United Nations, 1983).

Quantitative information on household food use is obtained by recall from the person identified as most responsible for food planning and preparation. Only expenditure data are collected for meals and snacks eaten away from home (Burk and Pao, 1976).

The nutritive value of the food used from household food supplies is calculated for the edible portion of food as brought into the household. Only the vitamin values are adjusted
for cooking loss. Thus, the nutritive value of household food supplies includes values not only of food eaten but edible food discarded in kitchen and at table. Therefore, the calculation procedures will overestimate the food energy and nutrient levels of foods eaten in many households. Data on household availability do not include information on consumption of foods or nutrients by individuals within the households.

Use of food at the household level has also been estimated by examination of items discarded as garbage (Gallo, 1980). Although this method has limited application with the types of studies considered in this report, the studies suggest a substantial amount of waste from foods brought into households.

C. ADEQUACY OF FOOD COMPOSITION DATABASES

Analysis and interpretation of information on both dietary intake and food availability require use of databases for food constituents. Such databases have been developed in both the public and private sectors. Extensive compilations of analytical values for nutrients are publicly available from Agriculture Handbook No. 8 (U.S. Department of Agriculture, 1976, 1977, 1978, 1979a,b, 1980a,b, 1982a,b, 1983, 1984a,b, 1986a,b). Content of 11 essential minerals in 234 commonly consumed foods is directly analyzed each year in the Selected Minerals in Foods Survey of the Total Diet Study conducted by FDA (Pennington et al., 1986). Fewer data are available for nonnutrient food constituents, food additives and GRAS ingredients, or contaminants such as pesticide residues, industrial chemicals, or radionuclides. Stewart (1983) has reviewed the state-of-knowledge regarding publicly available information on composition of foods, including nutrients, naturally occurring nonnutrients and toxicants, food ingredients, and inadvertent additives and contaminants of foods.

1. Databases for nutrients

Data in the food composition tables of Agriculture Handbook No. 8 are based on analytical values for samples of foods representing types of foods available on a nationwide, year-round basis (Exler, 1983). For example, data on vegetables may represent different growing years, growing areas, cultivars, processing techniques, lengths and conditions of storage, laboratories, and possibly different methods of analysis (U.S. Department of Agriculture, 1984a). In recent years, USDA has provided information on the standard error of reported food composition data (U.S. Department of Agriculture, 1976, 1977, 1978, 1979a,b, 1980a,b, 1982a,b, 1983, 1984a,b, 1986a,b). The standard error is not differentiated between methodologic error and sampling error. To date, information on reliability of estimates is
available for only a part of nutrient data and is not a part of the data banks used for surveys of dietary intake. Lack of appropriate reliability estimates remains a limitation of all currently available databases and may influence estimates of nutrient intakes to an unknown extent.

The quality of analytical values and the number of foods for which data are publicly available vary widely for specific nutrients (Stewart, 1983). Based only on quantity of data available, it would be reasonable to surmise that more reliable estimates could be made for intake of nutrients for which more analytical data are available. However, the state of development of analytical methods for nutrient content of foods varies considerably (Stewart, 1983). For example, assays may lack specificity for a nutrient, may be affected by other components in foods, or may not measure all active forms of a nutrient. Thus, the availability of analytical values does not automatically indicate that the data can be used to provide reliable estimates of nutrient intake. Sources of variability and bias in food composition tables and their effects on analysis of dietary intake data are discussed in the report of the Subcommittee on Criteria for Dietary Evaluation (National Research Council, 1986) and in Chapter IV of this report.

Values in the many nutrient data banks developed to meet the needs of specific studies may be based on analytical values reported by the U.S. Department of Agriculture, manufacturers' data, label statements, data compiled from the literature by the authors, unpublished analytical data, and compositions calculated from recipes or other special formulations (Hoover, 1984, 1985). However, much of the detail and documentation originally available from some of the tables is not incorporated into tables used for calculation of nutrient intakes in surveys. Conversely, even if precise data on food composition are available, their potential usefulness may be lost if identification or coding of food items are imprecise during the collection of dietary intake data.

Serious issues in any of the databases currently available include: validity of the data themselves; currentness of the data (updating of databases); and documentation of the data, including documentation of source, sampling reliability, and analytical method and reliability, and particularly, documentation of imputations. Specific examples of differences among data banks may include: 1) the number and description of food items listed and the amount of information available about commercially processed food products; 2) the nutrient composition values, specifically, source, vintage, completeness of nutrient information, analytical methodology, and quality control measures; and 3) the quantitative units used to express portion size and nutrient content (Jacobs et al., 1985; National
Research Council, 1986; Underwood, 1986). When data from different tables are combined to create data banks, gross errors may occur if differences between the tables are not reconciled (Underwood, 1986). In addition, comparison of intakes among studies using different data banks may be very misleading if analytical values differ substantially among the data banks.

Effects of the variance arising from food composition data and computation as an error source on estimates of intake are discussed in Chapter IV-G-3.

2. Databases for nonnutrient components of foods

Publicly available databases contain little information on nonnutrient components of foods. For some substances, analytical information is not available at all. For others, information is insufficient qualitatively or quantitatively to include in a database of food composition. In a few cases, analytical information may be available (for example, trypsin inhibitors in plant products), but it has not been organized for incorporation into a database of food composition.

Data on content of some pesticide residues, industrial chemicals, and radionuclides in 234 food items are available from the Total Diet Study of FDA. These food items are chemically analyzed each year for levels of more than 200 pesticide residues and industrial chemicals including lead, cadmium, and mercury and four radionuclides (cesium, strontium, iodine, and rubidium) (Gartrell et al., 1986a,b).
III. GENERAL STATISTICAL CONSIDERATIONS

General statistical principles and issues related to interpretation of data are presented in this chapter. Statistical issues specifically related to interpretation of dietary data will be discussed in Chapter IV.

A. SAMPLING

Survey or probability sampling is the discipline concerned with selection, observation, and analysis of samples from a population in order to make inferences about the whole population. An overview of conceptual issues in sampling is provided by Kish (1965).

The term "population" refers to a complete set of observable elements with defined characteristics. In the context of this report, the unit of observation is considered to be an individual or group of individuals observed one or more times. A population must be clearly delimited and may be defined by characteristics such as geographic area, time of observation, age, sex, socioeconomic status (SES), physiologic or disease state, or risk of an adverse outcome.

A sampling procedure selects a subset of population elements for measurement using rules fixed in probability theory. The sample on which classical statistical theory (Mood and Graybill, 1963) has been developed is the simple random sample,* in which each population element has an equal probability of selection into the sample independent of selection of each other element.

The factors most important for selecting a sample that best represents the population must be identified. If probability sampling is used, in the ideal case of sampling from the target population, mathematical propositions can be used to make inferences and generalizations. Judgment is always involved in deciding how closely the ideal case is approximated.

Sampled populations are rarely equivalent to the target population, even when sampling procedures follow formal sampling theory. For instance, in order to study the educational background of all nutritionists, one may have to settle for sampling from a list of members of a professional society. In this type

* The term "independent, identically distributed random variables" is used by mathematical statisticians to refer to this situation.
of situation, external information may provide a means for comparison of characteristics of society members with all nutritionists. Closeness of the sample to the target population is again a matter of judgment.

Use of complex sampling procedures will affect the precision of the observations. Stratification, in which the population is placed in subgroups and random selections made from each of the subgroups, tends to increase the precision of the observed data, compared to simple random sampling. Clustering, in which the population is placed in subgroups, the subgroups selected at random, and observations made only in the selected subgroups, tends to decrease precision. As discussed by Abraham (1986), use of complex sampling procedures requires the use of appropriate sample weighting factors to ensure correct interpretation of data.

Informal sampling includes haphazard or fortuitous samples such as samples of volunteer subjects, expert choice or representative sampling, and quota sampling. Great advances as well as definite errors have been made in many fields from studies not based on probability sampling. "No clear rule exists for deciding exactly when probability sampling is necessary, and what price should be paid for it. ... Probability sampling for randomization is not a dogma, but a strategy especially for large numbers" (Kish, 1965). Reference to a specific group of individuals is often important in descriptive studies. Even so, the importance of probability sampling may depend on the precision demanded. Casual observation may be sufficient to conclude that "everyone does it" versus "no one does it". Judgment plays a disproportionately larger role in informal sampling than in formal sampling.

Any sampling procedure divides a population into two parts: the sampled and the nonsampled elements. If the distributions of characteristics of interest are the same in both groups, then inferences from the sample group will hold for the entire population. There is potential for errors of inference to the extent that differences in distributions exist between the sample and the population.

The split into sampled and nonsampled elements is helpful in thinking about inferences in two other cases: nonresponse bias and clinical trials. The sample mean is not a perfect representation of the intended probability sample if there is nonresponse in the sample. In this case the mean of the (usually unknown) total probability sample is actually a weighted average of the (known) mean of the respondents and the (unknown) mean of the nonrespondents. To the extent that it is known that nonrespondents are similar to respondents, the researcher may be more confident in generalizing sample results to the entire population. Examination of characteristics of respondents and nonrespondents in studies in which
data are collected in a series of observations may aid in
determining the generalizability of data collected later in
the study because information about nonrespondents will be
available from data collected early in the study. This applies
to persons who refuse to participate at baseline as well as
to those who drop out of a study after several contacts. That
nonresponse effect may be quite important is suggested by the
observation that mortality was substantially higher in non-
participants than in participants in the initial random sam-
pling of the population of Framingham, Massachusetts for the
Framingham Heart Study (Gordon and Kannel, 1968).

In clinical trials, generalization occurs at two
levels. First, as in sampling in general, the participants
are drawn from a population. Results showing a difference
in effect between treatments may be generalized to the sampled
population based solely on mathematical propositions only in
the ideal case of probability sampling. This ideal is rarely
achieved because only those individuals willing to receive
treatment will participate in clinical trials. Second, a ran-
donm sample of all participants is selected for assignment to
treatment by a process of randomization, which is analogous
to random sampling. One or more groups receive the study treat-
ment(s) while one group receives the study control treatment.
Findings for each group are generalizable to the sample of all
participants. The principle of probability sampling forms the
mathematical basis for the comparability of the treated and
control groups in clinical trials. Generalization of results
from clinic populations to the total population must be made
with appropriate caveats (see Chapter III-I).

Data may contain useful information in the absence
of randomization or probability sampling (or both) and even
when biases are introduced by inequalities among comparison
groups. Often the geographic base or the willingness to par-
ticipate are not critical factors that determine outcome.
Consequently, special purpose sampling of volunteers with
specific characteristics may be the best option. Special
purpose sampling has often worked experientially; neverthe-
less, it has a weak mathematical basis.

The type of study and the question(s) asked determine
the type of sample that is needed. The following examples
illustrate this point. In a study of the physiologic effects
of coffee on endurance running (a study of a contrast or a
relationship), the most appropriate sample may be runners and
nonrunners who volunteer to submit to a specified protocol of
running and coffee/hot water drinking. However, to ascertain
how many runners are drinking coffee to increase endurance (a
study of an absolute level), it may be important to study a
probability sample that represents the population of interest.
Contrasts or relationships will often agree in the sampled and
nonsampled groups even when the absolute levels differ between those groups. For this reason, a strict mathematical base for sampling may be less critical for examination of health relationships than for examination of levels of health.

Data should be interpreted with a caution appropriate to the flaws in the data collection design. Findings should be replicable in independent studies when data are collected in a flawed design; thus, independent corroboration of findings is of great value.

B. BIAS AND RANDOM VARIATION IN MEASUREMENT

Evaluation of data should include not only consideration of the relationship of a sample to its parent population but also consideration of biases that may have been introduced by methods of selection, measurement, or analysis. Empirical information contained in the database should reflect as little as possible the preconceptions of the investigators (Gould, 1981).

As indicated in Chapter III-A, complex sampling designs may be viewed as variants of simple random sampling. Equal probability sampling methods tend not to introduce bias in inferences about the population parameters. However, methods such as sampling at unequal rates in different strata, or sampling elements with unequal probabilities do introduce bias that must be corrected by using appropriate stratum or case weights. The stratified sampling strategies used in national surveys such as the National Health and Nutrition Examination Survey (NHANES) and the Nationwide Food Consumption Survey (NFCS) necessitate use of weighting procedures to correct the resultant bias. Abraham (1986) has discussed use of weighting factors with NHANES data. Cluster sampling in clusters of unequal size tends to introduce bias in estimation of means since both the sample sum and the sample size are random variables. Kish (1965) has described the conditions under which this bias is acceptably low.

The pattern of variation in data can be estimated empirically. The use of the mean and standard deviation to describe normal variates and variability about this typical value is justified by reference to the normal distribution. Such interpretation of the probability structure of a variable from its mean and standard deviation is inadequate for nonnormal variates, such as consumption of alcoholic beverages for which a significant proportion of the population are non-drinkers and the amount of alcohol consumed per day among drinkers approximates a normal distribution. Even in cases of great skewness, bimodality and other nonnormal distributions, reference of mean and standard deviation to the normal
standard may provide some information; however, in this case, the extent of deviation from the normal standard should also be considered.

Whether to treat the population elements as if all are ruled by the same probabilities is a matter of judgment. If the underlying factors are randomly distributed, the probability structure may appropriately be regarded as identical between persons. However, two groups of persons may differ on some of these underlying factors and distributions may differ between the subgroups. Similarly, even with identical probability structures, the population elements may not be independently distributed. In such situations, the estimate of the variance of the mean will be incorrect.

The situation of correlated population elements arises most commonly from repeated measurements or cluster sampling. When two or more populations are pooled and treated as a single group, the mean estimated from such actually differing populations is biased. This situation, termed confounding, is common. Deconfounding by direct standardization is one straightforward method for removing this bias. Other standardization procedures are similar conceptually.

Repeated measurement within an observation unit induces correlations between observations, for example, collection of dietary data on consecutive days. In this situation, the sampled population must be regarded as a collection of elements observed repeatedly, thereby having a time dimension. Repeated measures may be either positively or negatively correlated.

C. COMPUTATION OF VARIANCES IN COMPLEX SAMPLING DESIGNS

Nonindependence of the distribution of population elements is encountered in some types of complex sampling designs. Generally, stratification does not introduce interdependence of elements, but multistage designs that include clustering at one or more stages do introduce interdependence. Clustering is used for logistic reasons in surveys such as the NHANES. Interdependence is also present in studies in which the individual is not the unit of analysis, for instance, in intervention studies using a school, worksite, physician practice, county, city, or state as the intervention unit. In prospective (cohort) studies of incidence of disease, contrasting subgroups often are formed through self-selection, such as through use of alcohol or cigarettes. Such persons have commonalities within groups that are not shared between groups. Intraclass correlation coefficients may be positive in self-selected subgroups. To the extent that intraclass correlation coefficients are positive, the subgroup forms a cluster (often from an unknown population), and the variance
of the mean as computed by \( s^2/n \) is too small. False positive hypothesis tests are one result; nonreplicability of studies is another. When significance tests are considerably less than \( p = 0.05 \), such false positive tests and nonreplicability of findings in other settings are less likely. This issue has generally not been addressed in the analysis of epidemiologic data.

Estimation generalizable to a specific human population is very important for descriptive work such as indications of nutrient status in the population at large or for some specific age-sex-ethnicity-SES group. The work of the U.S. Census Bureau, National Center for Health Statistics (NCHS), U.S. Department of Agriculture, and many other governmental agencies is particularly pertinent here. Many publications in the Vital and Health Statistics, Series 2 of NCHS (Feinleib and Feldman, 1984) contain specific examples of the effects of clustering. Variances appropriate to the sampling design, which represent the variability of the mean reasonably well (as judged by actual replication of the same mean in different experimental or observation settings), are necessary to avoid overly narrow confidence intervals and false positive hypothesis tests.

D. INTRAINDIVIDUAL VARIATION

The assumption of independent and identically distributed random variables (as in simple random sampling) fails to hold when the variance itself is not a unitary whole but consists of components. Thus, the concept of total variance arises, consisting in the simplest case of a component relating to variation within the individual unit under study (i.e., within a person across occasions) plus a component relating to variation of the different individual units about the mean of the distribution of all individuals (i.e., among persons). Intraindividual variation arises in many biological variables that change from day to day even when diet and other determining factors are constant (Jacobs et al., 1979; Liu et al., 1982). As discussed in Chapter IV, observed regressions and correlations are altered from the true values when intraindividual variation exists. This alteration represents a major problem in the study and interpretation of dietary and related complex behavioral data.

It is not known whether and to what extent intraindividual variability differs from person-to-person, nor is it known how statistically robust analyses are under the assumption that the level of intraindividual variability is constant across persons. At present, the assumption of constant intraindividual variability is warranted; however, the universality of this assumption requires further research.
E. IMPUTATION OF MISSING VALUES

Data collection is rarely complete. In prospective studies, disease status may be unknown or examinations may be missed entirely. In cross-sectional surveys, questionnaire items may be left blank through oversight, refusal to answer, or failure to understand the question; blood samples may be insufficient for all planned analyses; a participant may have to leave a survey center early; or an absurd answer may be given. Particular care must be taken in the case of potentially absurd values, i.e., an outlier caused by an error or merely a value that does not fit the analyst's preconceptions. Omission or imputation of values that do not fit preconceptions tend to make the data fulfill those preconceptions and should, therefore, be avoided.

Databases may have diminished value if a substantial amount of data is missing; in extreme situations a large amount of missing data may make a database unusable. Imputation for missing values can correct minor imbalances in the data and maintain cases in multivariate analyses which require that no values be missing. However, it must be realized that imputation is to some extent "making up data" and can be carried too far. Use of imputed values and procedures used for imputing missing values should always be documented in order to avoid misleading interpretation of data.

A common treatment for missing data is substitution of an average computed over the remainder of the database. This conservative approach reduces the variability in the data set and tends to reconfirm the values in the data. A variation to this approach which maintains overall variability is based on imputation with a study-wide average perturbed by addition of a random amount.

The next level of complexity takes into account characteristics of participants that are related to the missing value. A missing value may be replaced with an age-race-sex specific mean rather than a study-wide mean. The regression of a missing factor on several other factors may be calculated and a value predicted from information specific to a subject's missing value(s). Subgroup- or variable-specific methods tend to reflect results already in the data to a lesser extent than replacement of missing values with study-wide means.

Variable-specific predictions may be used in longitudinal studies when a value is missing at one time point in a series of observations. In such situations, the missing value is interpolated from the "before" and "after" values. Another strategy in longitudinal studies is to use the baseline value for subsequent missing values, based on the rationale that subjects who miss visits are not likely to have
complied with the experimental protocol. Obviously, this would be erroneous if the baseline value were expected to change over time, for example, body weight in growing children.

A special case for imputation in the analyses of dietary data involves the use of food composition tables. It is clearly inaccurate to use a zero value for a missing component. Simple imputation based on content of the substance in related foods is better than using zero. Food composition tables vary in their completeness with regard to analyzed values for particular constituents. For example, a table might not contain values for a certain nutrient in 50% of food items that contain significant amounts of the nutrient. At such a level of incompleteness, that particular food table should be considered inadequate for estimation of intake of that nutrient.

The impact of missing values may be assessed in certain situations. In longitudinal studies, characteristics of persons seen at entry but not at the follow-up visit can be compared to those who were seen at both times. A judgment can then be made about the extent of nonrepresentativeness created by loss to follow-up. By analogy, nutrient levels could be computed for each individual using only foods that do or do not have complete nutrient information. In regression analysis, missing values may be retained in analysis or a regression indicator variable may be used to assess the predictive value of missing data. This procedure is reassuring when missing data have no predictive value but causes interpretive difficulty when missing data do have predictive value. Such interpretive difficulty is an indication that missing data may be a problem.

Although dealing with missing information can be extremely complex and risky, data should be imputed in several situations. In multivariate analysis, some procedures that require matrix inversion may discard any cases that do not have complete data. Imputation may make it possible to enter into the analysis many cases that have the great majority of data available, particularly if data are missing in less critical variables. In nutrient analysis, use of weak estimates is preferable to zero values. In longitudinal studies in which follow-up is spotty, cases can sometimes be maintained in a trend line if only a few values are imputed, usually with the relatively strong method of interpolation. Finally, some statistical procedures require balanced data and can only be carried out if some missing data are imputed.

An alternative method to imputation which is under development involves analysis of the properties of the likelihood function of only observed data values (Laird and Ware, 1982; Louis, 1982). Although it is mathematically complex, this method has promise, particularly for analysis of repeated measures.
F. TRANSFORMATIONS AND UNUSUAL DISTRIBUTIONS

Variables do not always fit assumptions regarding normal distributions underlying most statistical procedures. Typically such variables are skewed heavily towards large values (serum triglycerides or energy expended in leisure time physical activity) or towards zero values (cigarette smoking). Data may be mathematically transformed to reduce the problem of distribution, for instance, by converting each value to its logarithm. Transformations selected are generally monotonic; that is, they preserve the ranking of the data across participants.

If the actual interval is important, transformations may alter the meaning in the data. For instance, the expenditure of 3000 kcal/week for physical activity is 1.5 times greater than 2000 kcal/week; the ratio of the natural logarithms is 8.0/7.6 = 1.05. For 300 and 200 kcal/week, the ratio of expenditure is also 1.5 and the ratio of the natural logarithms is 5.7/5.3 = 1.08. For 30 and 20 kcal/week, the ratio of expenditure is again 1.5, with the ratio of the natural logarithms 3.4/3.0 = 1.13. The actual interval in these three cases is 1000, 100, and 10 kcal/week, respectively and 0.4 in each case on the logarithmic scale. As ratios, these widely different energy expenditures are equal, with a tendency towards more weight at the low end of the log scale. Thus, the importance of large increases in physical activity expenditure will be minimized if the initial value is already a high level of activity.

In the case of cigarette smoking, the distribution is skewed but transformations are not appropriate. The zero value for persons who have never smoked is qualitatively different in some respects from any positive value. Smoking data should be analyzed with consideration for smoking history as well as current number of cigarettes smoked.

Despite these disadvantages, transformations that keep data in the same rank order allow statistical testing to be performed accurately. Fortunately, most linear regression procedures are rather insensitive to the assumption of normality from which they were derived and few errors should result if data are not transformed.

G. THE EFFECT OF MULTICOLLINEARITY ON STATISTICAL ESTIMATION

Large intercorrelations exist among energy intake and intake of specific nutrients and food groups. Nutrients are eaten in "packages"; certain nutrients often occur together in the same food items (Gordon et al., 1984). McGee et al.
(1984) reported a correlation of 0.77 between protein and fat intakes in 7088 men in the Honolulu Heart Study. High correlations were also reported for potassium intake with intakes of calcium, protein, and fat in the same study (Reed et al., 1985). Such high correlations indicate that intakes of the nutrients are not unrelated. Interpretation is correspondingly difficult.

An effect of multicollinearity involving matrix inversion is important statistically, particularly in the estimation of regression parameters. The mathematical situation in multicollinearity is that one row of a matrix is a linear function of another row. When multicollinearity approximately exists, as is the case when correlations on the order of $\geq 0.8$ exist in a data set, matrix inversion involves dividing by very small numbers. Inversion may be numerically possible, but small changes in the data set can result in dramatically different results. In this sense, statistical estimation is not stable. McGee et al. (1984) have explored all possible multiple logistic regressions of the association of coronary disease incidence in 10 years with energy, fat, protein, and carbohydrate intake. Different results were obtained, depending on the path through the regressions (deconfounding, forward stepping, or backward stepping). These differences arise because of the instability of statistical estimation in the presence of high correlations.

Statistical procedures cannot solve problems of multicollinearity. By necessity, investigators attempting to make etiologic attributions in the presence of multicollinearity employ mechanistic studies. Epidemiological studies are helpful in establishing the direction of an investigation but may have to be interpreted (as in the case of the work of McGee et al., 1984) as indicating either presence of substantial fat and protein intake or absence of dietary carbohydrate without coming to a conclusion about the active factor.

H. FACTOR ANALYSIS METHODS

Population-based studies generally are not the best means of studying the differential biologic action of individual nutrients, foods, or food groups because multicollinearity among nutrients prevents separation of effects of individual nutrients. An example to illustrate this point is the consideration of two nutrients with a correlation of 0.8. If these two variables are placed in decile categories and cross-tabulated, there will be almost no value in this table that deviates from the diagonal. Therefore, any data set representing food intake of free-living populations will carry little information separating the biologic actions of some pairs of nutrients.
Factor analysis is a statistical procedure that provides a means to recognize and highlight this clustering of variables and this ambiguity. Factors are weighted combinations of individual nutrients. These combinations are formed in such a way as to reflect the multicollinearity in the data and may be used in prediction of disease or other health relationships, just as food components or food groups would be. In dietary data, the components of the factors are biologically well-defined and the factors merely indicate how the several components coexist in diets. A major difficulty is biological interpretation of the factors. Little work has been done on this approach and it remains to be seen whether factor analysis will provide new insights into dietary relationships.

I. EPIDEMIOLOGICAL STUDIES AND INTERVENTION TRIALS

In addition to studies in which description of diet is the primary interest, dietary data are used extensively in epidemiological studies and intervention trials in an attempt to relate exposure to food components to health status or disease incidence. Considerations discussed in other sections of Chapter III and in Chapters II and IV also apply to the evaluation and interpretation of dietary data for epidemiological studies and clinical trials. Factors central to appropriate design and interpretation of epidemiological studies and intervention trials, including those having a dietary component, are summarized in Table 2. These issues are considered only briefly in this report; they are discussed comprehensively by Kleinbaum et al. (1982), Lee (1980), Schlesselman (1982), and Shapiro and Louis (1983).

1. Epidemiological studies

Interpretation of epidemiological data on relationships between diet and disease must be made cautiously because neither health status nor nutritional status can be observed directly and inferences about both must be made from indicators with a large degree of uncertainty (Ware et al., 1981). However, when coupled with clinical, pathological, and experimental evidence, results from epidemiological studies can provide a basis for making judgments of causality. The Expert Panel considers that the U.S. Surgeon General's Advisory Committee on Smoking and Health (1964) best summarizes the role of statistics in epidemiological studies, including its criteria for establishing causality:

Statistical methods cannot establish proof of a causal relationship in an association. The causal significance of an association is a matter of judgment which goes beyond any
### Table 2. Considerations for Design and Interpretation of Epidemiological Studies and Clinical Trials.

<table>
<thead>
<tr>
<th>TYPE OF STUDY</th>
<th>PURPOSES</th>
<th>PARAMETERS REQUIRING CLEAR DEFINITION</th>
<th>SAMPLING DESIGN</th>
</tr>
</thead>
<tbody>
<tr>
<td>EPIDEMIOLOGICAL (OBSERVATIONAL)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cross-sectional Study</td>
<td>● To estimate a prevalence rate (e.g., for levels of intake of food components or for diseases)</td>
<td>● Study objectives (hypotheses or research questions)</td>
<td>● Random sample selected from the target population, (e.g., NHANES and NFCS)</td>
</tr>
<tr>
<td></td>
<td>● To estimate average levels of risk factors</td>
<td>● Target population</td>
<td>● Nonrandom sample (e.g., employee population, community study)</td>
</tr>
<tr>
<td></td>
<td>● To survey certain issues for policy</td>
<td>● Endpoints</td>
<td></td>
</tr>
<tr>
<td></td>
<td>● To study associations between two factors</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cohort Study (Prospective)</td>
<td>● To study the relationships between potential risk factors and a particular outcome</td>
<td>● Study objectives</td>
<td>Usually not a random sample (e.g., Framingham and Western Electric studies)</td>
</tr>
<tr>
<td></td>
<td>● To study the relative frequencies of an outcome in people with or without some exposure factor</td>
<td>● The outcome under study and the exposure of interest</td>
<td>● Single cohort: a sample is selected at baseline and then followed over time</td>
</tr>
<tr>
<td></td>
<td></td>
<td>● Target population</td>
<td>● Multiple cohorts: several cohorts which have different exposure status are selected at baseline and then followed over time</td>
</tr>
<tr>
<td></td>
<td></td>
<td>● Endpoints</td>
<td>● Historical cohort: all outcomes have occurred before the start of the investigation; the cohort is established and their experience assessed from existing records</td>
</tr>
<tr>
<td>Case-control Study (Retrospective)</td>
<td>● To test etiologic hypotheses for specific rare outcomes or diseases</td>
<td>● Study objectives</td>
<td>● Random sampling for cases and controls</td>
</tr>
<tr>
<td></td>
<td></td>
<td>● The outcome under study and the exposure of interest</td>
<td>- total random sampling</td>
</tr>
<tr>
<td></td>
<td></td>
<td>● Eligibility of cases and controls</td>
<td>- stratified random sampling</td>
</tr>
<tr>
<td></td>
<td></td>
<td>● Sources of cases -- hospital or disease registry</td>
<td>- Matched sampling of cases with controls</td>
</tr>
<tr>
<td></td>
<td></td>
<td>● Sources of controls -- hospital, clinic, and community</td>
<td>- reduce biases</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>- difficult to determine matching criteria</td>
</tr>
<tr>
<td>INTERVENTION STUDIES (Clinical Trials)</td>
<td>● To determine the relative merits of the treatments</td>
<td>● Study objectives</td>
<td>Nonrandom sample</td>
</tr>
<tr>
<td></td>
<td></td>
<td>● Target population</td>
<td>● Two or more independent groups design</td>
</tr>
<tr>
<td></td>
<td></td>
<td>● Control groups</td>
<td>● Sequential design</td>
</tr>
<tr>
<td></td>
<td></td>
<td>● Inclusion and exclusion criteria</td>
<td>● Randomization of subject assignment to groups</td>
</tr>
<tr>
<td></td>
<td></td>
<td>● Treatment protocol</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>● Endpoints</td>
<td></td>
</tr>
<tr>
<td>DATA ANALYSES</td>
<td>COMMENTS</td>
<td></td>
<td></td>
</tr>
<tr>
<td>---------------</td>
<td>----------</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
| Methods: analysis of variance (ANOVA), analysis of covariance, correlation (partial correlation) analyses, linear regression, and risk ratio analyses | - Cross-sectional studies are useful for generating new hypotheses.  
- Cross-sectional studies are usually easier to conduct and more feasible than cohort or case-control studies; however, they provide no information on temporal sequences.  
- Nonrandom sampling limits the generalizability of the results.  
- Nonrandom samples may be useful for studying association between two factors; however, they are not recommended for the first three purposes. |
| Strategies:  
- compare the baseline characteristics of the subgroups of interest  
- adjust the confounding variables | - Cohort studies can provide information on temporal sequences.  
- For single and multiple cohort:  
  - the study examines the incidence (new cases) of disease if the outcome status is determined at baseline.  
  - the study examines the mixture of incidence and prevalence of the disease if the outcome status is not determined at baseline.  
- Nonrandom sampling limits the generalizability of the results.  
- Large losses to follow-up (nonresponse) may bias results.  
- Cohort studies allow examination of multiple potential effects of a given exposure; however, study factors may change over time, making findings irrelevant. |
| Methods: risk ratio (relative risk and odds ratio) analyses, logistic regression, analysis of covariance, multiple comparison, Cox regression | - Case-control studies are appropriate for studying rare outcomes or diseases.  
- Case-control studies require fewer subjects than cohort studies.  
- Data collected by recall may be inaccurate; validation of data is difficult.  
- Current dietary status may not reflect predisease status.  
- Selection of appropriate controls may be difficult.  
- Control for confounding variables may be incomplete.  
- Case-control studies provide no information on temporal sequences. |
| Methods: odds ratio to approximate relative risk, odds ratio for matched pairs (one-to-one match and one-to-n match), logistic regression for case-control study (unmatched), and logistic regression for matched case-control study | - Target population is not the entire U.S. population.  
- Nonrandom sampling limits the generalizability of results.  
- Randomization of subject assignment to groups can balance for known and unknown (confounding) factors. |
| Strategies: may differ depending on comparability of the groups (drop out)  
Methods: rate difference, life table analysis, Kaplan Meir survival analysis (product limit), Cox regression, Wilcoxon test, logrank test, two sample t-test, ANOVA, analysis of covariance, and multiple comparisons |
statement of statistical probability. To judge or evaluate the causal significance of the association between the attribute or agent and the disease, or effect upon health, a number of criteria must be utilized, no one of which is an allsufficient basis for judgment. These criteria include:

a) The consistency of the association
b) The strength of the association
c) The specificity of the association
d) The temporal relationship of the association
e) The coherence of the association.

2. Intervention trials

Data from intervention trials or clinical studies having a dietary component also must be interpreted with caution. Blinding of subjects (and, in some cases, investigators) to the treatments or interventions applied cannot be achieved in studies in which total dietary patterns are altered (for example, the influence of the level of dietary fat on development of certain forms of cancer). Gaining compliance in studies involving modification of the total diet may be difficult, resulting in incomplete penetration of the treatment or intervention. Also, subjects may report what they think the investigator wants to hear. There is often little means to determine adherence to the protocol and an inaccurate estimate of a subject's response to the treatment or intervention may result. Such misclassification of individuals lowers correlation and regression coefficients and underestimates associations between diet and health status (see Chapter IV-C).

Additionally, generalizability of results from clinical trials or intervention studies to the target population may be limited because, in most cases, the sampled population is not the same as the target population (see Chapter III-A).
IV. ISSUES RELATED TO STATISTICAL ANALYSIS OF DIETARY INTAKE DATA

For studies involving dietary surveys, the underlying assumption is that people have maintained or will maintain their dietary habits during a target period of time. Case-control studies on the relationship between a dietary factor or factors and a disease assume that the disease cases maintained dietary habits acquired before the onset of the disease. The study can then test the hypothesis that the dietary factors of interest were different in the habitual diets of the case and control subjects. Cross-sectional studies generally assume that the current dietary habits have been unchanged for a long period of time, and cohort studies further assume that the dietary habits will be continued in the future. In other words, during the time period of interest, there are no mean changes in each of the dietary factors and the day-to-day variation of each dietary factor is just the random variation around its mean. For most studies, this basic assumption is made implicitly. In some cases, the assumption may not even be reasonable because measurable changes in dietary habits have occurred.

The impact of dietary intake methodology on statistical analysis and interpretation must be considered within the framework that this assumption is reasonable. How well has the implied average intake been estimated? What is the impact of any error in this estimate on subsequent statistical analyses?

A. EFFECT OF METHODOLOGY ON STATISTICAL PARAMETERS

Use of specific types of dietary intake methodology for estimation of specific parameters was discussed in Chapter II-A-2. The most salient points of that discussion are summarized in the following paragraphs.

Although actual intake data obtained by quantitative daily consumption methods are relatively more accurate than data collected by food frequency methods, problems arise when the number of records or recalls is inadequate to assess an individual's habitual intake. The major difficulty in analysis for quantitative daily consumption data is the large amount of intraindividual variability in daily intake. Some sources of intraindividual variation are discussed in Chapter IV-G.

Additionally, minor weaknesses are associated with data collected by quantitative daily consumption methods. These methods tend to miss foods that are infrequently consumed. Unless the study is designed properly, the data collected will not take into account variation such as day of the week (weekday or weekend) or season of the year. Although study design may lessen the effects of these problems, a large number of repeated food records and/or recalls may be difficult to collect.
The theoretically correct way to estimate the average dietary intake of an individual is to randomly select a large number of days over the target period of time and then to calculate the mean intake of the dietary factors on these days. However, most studies have used the intake of only 1 day or a few days to estimate the mean intake. As a consequence, the large intraindividual variation and the inadequate number of daily measurements affect the ability to detect the potential relationships between dietary factors and biological risk factors or diseases.

Frequency methods directly estimate average intake of foods or food components during the target period; therefore, intraindividual variation is not a serious problem. Collection of data by frequency questionnaires is generally easier than collection of a large number of food records or recalls; however, the data are less accurate. Combining of categories of food frequency in analysis causes difficulties in making statistical inferences because arbitrary combinations may artificially inflate or deflate estimates of intake.

Because it is very difficult for subjects to recall exact foods eaten, meal by meal or day by day, even in the recent past, food frequency questionnaires generally ask about the pattern of foods consumed rather than description of meals eaten. Therefore, description of actual intake is imprecise. Two possible types of error result: systematic bias in reporting of frequency (over- or under-reporting of total food intake or of intake of certain classes of foods) and inaccuracy of the estimation of portion sizes. Depending upon instrument design, there may be errors also in the identification of specific food items (i.e., the pooling of similar items that are not similar in all compositional characteristics). Despite these problems, food frequency methods can be a useful tool. If a frequency method can preserve the rank order of each dietary factor in the population, the statistical analyses for examining relationships may still be valid.

An inadequate number of daily consumption records and an inaccurate food frequency method may both produce serious measurement errors for the corresponding data; i.e., the estimated intake of dietary factors may be very different from the true mean intake. Consequently, the association of dietary factors with a disease, for example, may be attenuated. The pattern of the attenuation varies from analysis to analysis. The remainder of this chapter summarizes the impact of measurement errors on specific types and aspects of statistical analyses.
B. UNIVARIATE ANALYSIS

For quantitative daily consumption methods, the presence of intraindividual variation increases the total observed variance in intake, which is actually a summation of the intraindividual variance and the interindividual variance. That is, the observed variance is greater than the real variance of the usual intake in the population. Consequently, the distribution curve of the observed intake is "flattened."

For estimating the mean and median of usual intake, the enlarged variance resulting from the presence of intraindividual variation does not bias the point estimation; however, it substantially increases the widths of the confidence intervals. This problem may not be very serious if the sample size is large. However, estimation of centiles can be biased by the enlarged variance. As a consequence of the flattened distribution curve, the higher centiles (i.e., greater than the median) and the lower centiles (i.e., less than the median) are overestimated and underestimated, respectively. Similarly, the proportion of participants with an intake above a cutoff point to the right of the median, or below a cutoff point to the left of the median, is inflated. The prevalence of "low" or "high" intake is, therefore, likely to be overestimated. Because intraindividual variation tends to equal or exceed the interindividual variation for most nutrients in free-living populations, this problem could be quite grave. Table 3 provides the ratio of the observed variance of usual intake to the interindividual variance corresponding to different sizes of intraindividual variance and different numbers of measurements. The square roots of the values in Table 3 represent a multiplicative factor for increase in the confidence interval for the population mean intake. It can be seen from this table that one way to reduce the "inflation" of the observed variance is to increase the number of measurements. For example, when the ratio of intraindividual to interindividual variance is 3, the variance of the observed value generated from one measurement is four times the interindividual variance. However, the variance of the observed values generated from seven measurements may be reduced to 1.43 times the interindividual variance. Data sets based on observations covering several days will present less bias in the estimate of centiles or prevalence of "high" or "low" intake than data sets based on a single day's observations.

The Subcommittee on Criteria for Dietary Evaluation (National Research Council, 1986) illustrated an alternate way to treat the problem with even single replicates (two measurements of 1-day intakes) or a statistically adequate sample of replicates to estimate the partitioning of variance between the intraindividual and interindividual components under an assumption that intraindividual variation is constant across persons.
Table 3. Ratios\(^1\) of the Observed Variances\(^2\) to Interindividual Variances.

<table>
<thead>
<tr>
<th>(\frac{\sigma^2}{\sigma r} )</th>
<th>Number of Measurements/Person</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>0.00</td>
<td>1.00</td>
</tr>
<tr>
<td>0.25</td>
<td>1.25</td>
</tr>
<tr>
<td>0.50</td>
<td>1.50</td>
</tr>
<tr>
<td>1.00</td>
<td>2.00</td>
</tr>
<tr>
<td>1.50</td>
<td>2.50</td>
</tr>
<tr>
<td>2.00</td>
<td>3.00</td>
</tr>
<tr>
<td>2.50</td>
<td>3.50</td>
</tr>
<tr>
<td>3.00</td>
<td>4.00</td>
</tr>
<tr>
<td>3.50</td>
<td>4.50</td>
</tr>
<tr>
<td>4.00</td>
<td>5.00</td>
</tr>
<tr>
<td>4.50</td>
<td>5.50</td>
</tr>
<tr>
<td>5.00</td>
<td>6.00</td>
</tr>
</tbody>
</table>

\(^1\) Tabulated values calculated by Panel members from the formula \(1 + \frac{\sigma_a^2}{(k)\sigma_r^2}\) where \(\sigma_a^2\) = intraindividual variation, \(\sigma_r^2\) = interindividual variation, and \(k\) = number of measurements.

\(^2\) The variance of the average value of the corresponding number of measurements.

* Ratio of intraindividual to interindividual variances.
It is preferable that the replicated measures be made on independent days to reduce any effect of autocorrelation between intakes on adjacent days, and that they be designed to incorporate the effect of known sources of variance. The observed distribution can then be adjusted. To do this, the Subcommittee suggested that the data set be normalized by appropriate data transformation and then an analysis of variance (ANOVA) be run to estimate the interindividual variance. In turn this estimate can be used to adjust the observed distribution on the simple basis that for any point on the observed distribution \( x \), the new value \( x' \) would be:

\[
x' = \mu + (x - \mu) \cdot \frac{SD_{interindividual}}{SD_{observed}}
\]

where \( \mu \) = the group mean and \( SD \) = standard deviation (observed and interindividual). Following this adjustment, the data would be retransformed to give an adjusted distribution in the original units (National Research Council, 1986). Often, but not always, a logarithmic transformation followed by an exponential transformation will be satisfactory. The most appropriate transformation to yield the best fit to the normal distribution can be determined by statistical techniques.

This adjusted distribution can be used in the estimation of centiles and the prevalence of "low" or "high" intakes. The bias may be greatly reduced. This method is not recommended for studies with very small sample sizes of replicate samples because the estimate of the variance of usual intakes may be very inaccurate in such cases.

The Subcommittee on Criteria for Dietary Evaluation (National Research Council, 1986) applied simulation and theoretical analyses to demonstrate that this procedure (or the use of the means of multiple days of observation) would also serve to reduce or remove the effect of other sources of random variation within individuals that might otherwise lead to a bias in a prevalence estimate. Such other sources of random error might include the implicit variation of food composition, random under- and over-reporting and random under- and over-recording (interviewer error), etc. However, the Subcommittee cautioned that adjustments of this type would not correct for any true bias in the data such as systematic under- or over-reporting, or systematic error in the food composition tables. Unless the replicate measures have been designed to include such effects as weekday vs. weekend differences, seasonal differences, etc., the approach will not correct for these (National Research Council, 1986).

In conclusion, as long as replicate days of intake are available and the size of the sample is very large, there are feasible approaches to the estimation of the distribution of usual intake for populations or population subgroups of
reasonable size. These techniques should be applied if the data represent intake for fewer than 3 days. It would be desirable to apply them even when the database represents 7 days of intake. The exception to this rule would be an analytical situation in which the question posed really did relate to a need to consider actual intake on any single day or group of days for which intake is measured directly. An example might arise from the study of an acute (single) exposure to a toxin rather than chronic intake. In any case, the statistical approach to adjustment and analysis must be derived from the biological question being asked.

Estimation of prevalence when the cutoff point is not the same for each individual is discussed in Chapter V-B and C.

For food frequency methods, the main problems involved in statistical analysis are the measurement error (error of the estimated value caused by the inaccuracy of the measuring instrument) and the possible systematic bias (over- or underestimation of intake). For most nutrients, the measurement error for a well-designed frequency method tends to be smaller than the intraindividual variance of 1-day quantitative daily consumption methods. The impact of the measurement error on the estimation of mean, median, centiles, and prevalence of "low" and "high" intake is similar to that of intraindividual variation. Systematic biases are still serious in some cases and repeat measures of the frequency data do not reduce these biases. In addition, the measurement error of a frequency method cannot be estimated directly. Thus, the use of a statistical correction for these biases is not generally possible. The Expert Panel recommends that data collected by any frequency method should be compared with data collected by multiple measurements of daily consumption methods randomly scheduled throughout the target period. In this way, the validity of the frequency method can be better ascertained.

C. BIVARIATE ANALYSIS

For bivariate analyses, the impact of the intraindividual variation varies from analysis to analysis. For simple correlations, the coefficient is attenuated by the factors tabulated in Table 4. The first column lists the ratio of intraindividual variances and the first row lists the number of measurements used to estimate the dietary intake. For example, if the ratio of intraindividual variances for a dietary factor is 3 and two 24-hour recalls are used to estimate the intake, the correlation coefficient between the estimated intake and some biological variable is 63% of the true correlation. The correlation coefficient between the estimated value generated from seven 24-hour recalls and some biological...
Table A. Attenuation Factors for Simple Correlation Coefficients.

<table>
<thead>
<tr>
<th>( \frac{\sigma_a^2}{\sigma_r^2} )</th>
<th>Number of Measurements/Person</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>1</td>
</tr>
<tr>
<td>0.00</td>
<td>1.00</td>
</tr>
<tr>
<td>0.25</td>
<td>0.89</td>
</tr>
<tr>
<td>0.50</td>
<td>0.82</td>
</tr>
<tr>
<td>1.00</td>
<td>0.71</td>
</tr>
<tr>
<td>1.50</td>
<td>0.63</td>
</tr>
<tr>
<td>2.00</td>
<td>0.58</td>
</tr>
<tr>
<td>2.50</td>
<td>0.53</td>
</tr>
<tr>
<td>3.00</td>
<td>0.50</td>
</tr>
<tr>
<td>3.50</td>
<td>0.47</td>
</tr>
<tr>
<td>4.00</td>
<td>0.45</td>
</tr>
<tr>
<td>4.50</td>
<td>0.43</td>
</tr>
<tr>
<td>5.00</td>
<td>0.41</td>
</tr>
</tbody>
</table>

1 Tabulated values calculated by Panel members from the formula \( \sqrt{1/1 + \left[ \frac{\sigma_a^2}{k \sigma_r^2} \right]} \) where \( \sigma_a^2 \) = intraindividual variation, \( \sigma_r^2 \) = interindividual variation, and \( k \) = number of measurements.

* Ratio of intra- to interindividual variances.
variable is 84% of the true correlation. Table 4 can also be used to calculate the number of measurements required for accurately characterizing an individual's dietary intake. For example, if the ratio of intra- to interindividual variance is 2.5, 10 to 12 measurements are necessary to limit the reduction of the correlation to be less than 10%. The statistical test for simple correlation is also affected by the intraindividual variation; however, if the sample size of the study is large and if the true correlation is not small, the test result will still be significant.

For bivariate analysis, the attenuation can be corrected statistically. Table 5 provides the multipliers to correct the attenuation for simple correlation. (Note: Table 5 also presents the inverse of the ratios calculated on p. 41.) When the ratio of intra- to interindividual variance is 3 and one measurement is used for estimating dietary intake, the correlation can be corrected by multiplying the correction factor 2.00 by the correlation coefficient. However, correction of the correlation has to be done with caution. When the sample size is not large, e.g., less than 100, the sampling error of both the correction factor and the correlation coefficient could be very large. Rosner and Willett (1987) have recently proposed a different approach to compute confidence intervals for the true correlation based on the observed simple correlation and the ratio of intra- to interindividual variances.

Table 6 presents the attenuation factors corresponding to different ratios of intra- to interindividual variance and different numbers of measurements for simple linear regression. When the ratio is equal to 2 and three measurements are used, the regression coefficient of a biological variable on the estimated value of the dietary factor is 60% of the true coefficient. The values in Table 3 are actually the correction factors for regression coefficients. The use of this table is similar to Table 4 for simple correlation analysis. Again, the correction has to be made with caution in order to avoid misleading analyses.

Attenuation is not very serious with respect to bias when the mean intake of a dietary factor is compared among groups. Although large intraindividual variation increases the residual sum of squares and thus reduces the power of the test, the mean value for each group is not changed by intraindividual variation. If the sample size is very large, the power is likely to be large (even with the attenuation) and the test results are not greatly affected.

When individuals are classified into different groups by a dietary variable, large intraindividual variation will cause a large proportion of individuals to be misclassified. The extent of misclassification of nutrient intakes of individuals into quintiles using 1-, 3-, or 7-day records or recalls
Table 5. Ratios\(^1\) of the Observed Standard Deviation\(^2\) to Interindividual Standard Deviation.

<table>
<thead>
<tr>
<th>(\frac{\sigma_a^2}{\sigma_r^2})</th>
<th>Number of Measurements/Person</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>----</td>
<td>---</td>
</tr>
<tr>
<td>0.00</td>
<td>1.00</td>
</tr>
<tr>
<td>0.25</td>
<td>1.12</td>
</tr>
<tr>
<td>0.50</td>
<td>1.22</td>
</tr>
<tr>
<td>1.00</td>
<td>1.41</td>
</tr>
<tr>
<td>1.50</td>
<td>1.58</td>
</tr>
<tr>
<td>2.00</td>
<td>1.73</td>
</tr>
<tr>
<td>2.50</td>
<td>1.87</td>
</tr>
<tr>
<td>3.00</td>
<td>2.00</td>
</tr>
<tr>
<td>3.50</td>
<td>2.12</td>
</tr>
<tr>
<td>4.00</td>
<td>2.24</td>
</tr>
<tr>
<td>4.50</td>
<td>2.35</td>
</tr>
<tr>
<td>5.00</td>
<td>2.45</td>
</tr>
</tbody>
</table>

\(^1\) Tabulated values calculated by Panel members from the formula \(\sqrt{1 + \frac{\sigma_a^2}{(k)\sigma_r^2}}\) where \(\sigma_a^2\) = intraindividual variation, \(\sigma_r^2\) = interindividual variation, and \(k\) = number of measurements.

\(^2\) The standard deviation of the average value of the corresponding number of measurements.

* Ratio of intra- to interindividual variances.
Table 6. Attenuation Factors\(^1\) for Simple Linear Regression Coefficients.

<table>
<thead>
<tr>
<th>$\frac{\sigma^2_a}{\sigma^2_r}$</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>10</th>
<th>12</th>
<th>14</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.00</td>
<td>1.00</td>
<td>1.00</td>
<td>1.00</td>
<td>1.00</td>
<td>1.00</td>
<td>1.00</td>
<td>1.00</td>
<td>1.00</td>
<td>1.00</td>
<td>1.00</td>
<td>1.00</td>
</tr>
<tr>
<td>0.25</td>
<td>0.80</td>
<td>0.89</td>
<td>0.92</td>
<td>0.94</td>
<td>0.95</td>
<td>0.96</td>
<td>0.97</td>
<td>0.97</td>
<td>0.98</td>
<td>0.98</td>
<td>0.98</td>
</tr>
<tr>
<td>0.50</td>
<td>0.67</td>
<td>0.80</td>
<td>0.86</td>
<td>0.89</td>
<td>0.91</td>
<td>0.92</td>
<td>0.93</td>
<td>0.94</td>
<td>0.95</td>
<td>0.96</td>
<td>0.97</td>
</tr>
<tr>
<td>1.00</td>
<td>0.50</td>
<td>0.67</td>
<td>0.75</td>
<td>0.80</td>
<td>0.83</td>
<td>0.86</td>
<td>0.88</td>
<td>0.89</td>
<td>0.91</td>
<td>0.92</td>
<td>0.93</td>
</tr>
<tr>
<td>1.50</td>
<td>0.40</td>
<td>0.57</td>
<td>0.67</td>
<td>0.73</td>
<td>0.77</td>
<td>0.80</td>
<td>0.82</td>
<td>0.84</td>
<td>0.87</td>
<td>0.89</td>
<td>0.90</td>
</tr>
<tr>
<td>2.00</td>
<td>0.33</td>
<td>0.50</td>
<td>0.60</td>
<td>0.67</td>
<td>0.71</td>
<td>0.75</td>
<td>0.78</td>
<td>0.80</td>
<td>0.83</td>
<td>0.86</td>
<td>0.88</td>
</tr>
<tr>
<td>2.50</td>
<td>0.29</td>
<td>0.44</td>
<td>0.55</td>
<td>0.62</td>
<td>0.67</td>
<td>0.71</td>
<td>0.74</td>
<td>0.76</td>
<td>0.80</td>
<td>0.83</td>
<td>0.85</td>
</tr>
<tr>
<td>3.00</td>
<td>0.25</td>
<td>0.40</td>
<td>0.50</td>
<td>0.57</td>
<td>0.63</td>
<td>0.67</td>
<td>0.70</td>
<td>0.73</td>
<td>0.77</td>
<td>0.80</td>
<td>0.82</td>
</tr>
<tr>
<td>3.50</td>
<td>0.22</td>
<td>0.36</td>
<td>0.46</td>
<td>0.53</td>
<td>0.59</td>
<td>0.63</td>
<td>0.67</td>
<td>0.70</td>
<td>0.74</td>
<td>0.77</td>
<td>0.80</td>
</tr>
<tr>
<td>4.00</td>
<td>0.20</td>
<td>0.33</td>
<td>0.43</td>
<td>0.50</td>
<td>0.56</td>
<td>0.60</td>
<td>0.64</td>
<td>0.67</td>
<td>0.71</td>
<td>0.75</td>
<td>0.78</td>
</tr>
<tr>
<td>4.50</td>
<td>0.18</td>
<td>0.31</td>
<td>0.40</td>
<td>0.47</td>
<td>0.53</td>
<td>0.57</td>
<td>0.61</td>
<td>0.64</td>
<td>0.69</td>
<td>0.73</td>
<td>0.76</td>
</tr>
<tr>
<td>5.00</td>
<td>0.17</td>
<td>0.29</td>
<td>0.38</td>
<td>0.44</td>
<td>0.50</td>
<td>0.55</td>
<td>0.58</td>
<td>0.62</td>
<td>0.67</td>
<td>0.71</td>
<td>0.74</td>
</tr>
</tbody>
</table>

\(^1\) Tabulated values calculated by Panel members from the formula $1/1 + \left[\frac{\sigma^2_a}{(k)\sigma^2_r}\right]$ where $\sigma^2_a = $ intraindividual variation, $\sigma^2_r = $ interindividual variation, and $k = $ number of measurements.

\* Ratio of intra- to interindividual variances.
is illustrated by recent work of Johnson (1986). Odds ratios calculated from misclassified data are biased toward unity and the decrease can be very large for large odds ratios (Johnson, 1986). As a consequence of misclassification, associations are attenuated and it becomes more difficult to detect the strength of association between diet and disease.

For ANOVA, the difference (main effect) in group means will be seriously reduced and the test will be greatly influenced by the misclassification (Liu, 1986). In these cases, increasing the number of replicated measurements is the only means for reducing the misclassification.

Based on these discussions, it can be concluded that the impact of misclassification on ANOVA may be serious. The ratio of the intra- to interindividual variances for a physiological variable (e.g., blood pressure, serum cholesterol, or plasma glucose) is generally smaller than that of a dietary variable. Therefore, to study the relationship between a physiological variable and a dietary variable, it is suggested that individuals be classified by the physiological variable with comparisons of the mean values of the dietary variable among the different groups.

The impact of the measurement error on bivariate analyses is very similar to that of the intraindividual variation but to a lesser degree. Again, there is no statistical correction for the attenuation and repeating the number of measurements will not reduce the attenuation. Furthermore, systematic bias does not have a serious impact on bivariate analyses. Systematic bias is not likely to attenuate the correlation coefficients in bivariate analyses but will yield an error in the intercept in linear regression.

D. MULTIVARIATE ANALYSIS

The impact of intraindividual variation on partial correlation, multiple linear regression, and multiple logistic regression is more complicated than its effect on simple correlation or regression (Liu, 1986). When the covariates (or the controlled variables) do not have intraindividual variation, the correlation or regression coefficient is always attenuated. The attenuation in these analyses is generally more serious than that in simple correlation or regression (Liu, 1987). The attenuation is very serious if a large proportion of the variation of the variable of interest can be explained by the covariates. When the covariates also have intraindividual variation, the coefficient may increase or decrease depending on many other factors (Kupper, 1984; Liu, 1987). For multivariate analyses, statistical correction of the attenuation is generally not feasible. The only way to reduce the attenuation of the potential biases is by increasing the number of measurements for each person.
It must be recalled that measurement error which still remains in the food frequency methods also impacts in exactly the same way as intraindividual variation discussed above. Again, the error in correlation coefficients cannot be reduced by increasing the number of replicate measurements or by statistical corrections. The effect of systematic bias of the method on bivariate analyses will not have a serious influence on these analyses. That is, it will not preclude detection of relationships although the description of the relationship (e.g., intercept in linear regression) will be in error.

E. POOLING DATA ACROSS STUDIES

Pooling of data across intervention studies has been examined as a means to strengthen the observations and increase the generalization of results from individual studies (Canner; 1983; Marmot, 1986; The Pooling Project Research Group, 1978; Rose, 1985). Pooling may be as simple as the observation that all or most studies are supportive of a point or as complex statistical procedures employed by Canner (1983). Pooling is based on the assumptions that a treatment or procedure will have the same effect under a variety of similar experimental conditions and that random variation in individual studies with small sample sizes may prevent the finding of conclusive outcomes. The pooled result should be less subject to random variation and be a more accurate reflection of the underlying situation.

The critical issue in pooling data is the degree of comparability of the methods employed and data obtained. Pooling can proceed under moderately different experimental or observational conditions. The Expert Panel considered minimum requirements for pooling to be: 1) a comparable experimental condition in the several studies and 2) a common theory linking this experimental condition to a common outcome measured in comparable ways. Factors to be considered in deciding whether these requirements are met include:

- study designs sufficiently similar to merge experimental groups across studies;
- similar study protocols, including intervention and measurement techniques, across studies;
- study populations sufficiently similar to allow pooling; and,
- questions and measurements sufficiently similar to consider that the same variables are being studied.
Statements about means and percentiles of dietary variables will be strongly affected by precision of method of data collection and particularly by intraindividual variation. Pooling of studies that use quantitative daily consumption methods over different numbers of days should be done carefully with appropriate weighting for differential variance. Similarly, pooling across different food frequency methods and pooling across food frequency and quantitative daily consumption methods may be problematic.

Bivariate relationships or multivariate relationships in which major intraindividual variation is associated with only the dietary variables will be attenuated less if more precise dietary measures are used. Therefore, pooling of relationship data that use different dietary data collection methods is also problematic. Statistical correction for attenuation may be made before pooling; however, the variability in deattenuation factors has not been studied and it is not known whether studies proposed for pooling should each be corrected in the same way or in different ways. Furthermore, imprecision in multiple variables in an analysis may strengthen or weaken measures of association, depending on the total configuration of within-person error. Statistical correction and, therefore, pooling is hazardous in such instances.

F. ASSESSMENT OF DIETARY TIME TRENDS

The assessment of time trends is closely related to the issue of pooling data; however, instead of two populations defined according to geography, ethnicity, disease status, or gender, the data considered for pooling are defined by the times at which they were collected. Dietary time trends may be assessed by following national food supply disappearance data over time, the same cohort over time, or by comparing prevalence data generated by studying representative samples of a population at different times.

1. General considerations

Time trend assessment begins with separate quantification of dietary intake or availability at each time point. Issues in interpretation of such data were discussed in earlier sections of this chapter; all points made in those sections apply equally to time trend data. Several additional points should also be noted.

At best, study populations at each time point should be defined identically, sampling procedures should be equivalent, and measurement procedures should be identical. Such stringent conditions are rarely met. Changes over time may occur because
the target population has changed over time. The target population may consist of the same individuals whose characteristics change with time or it may consist of different individuals at different times. Methodologic changes also have occurred over time, for example, in the estimation of national disappearance of fats and oils (Call and Sánchez, 1967). Discrepancies are often found in comparisons of intake estimates based on disappearance data and individual intake data (Food and Agriculture Organization of the United Nations, 1983).

National disappearance data supply the only database for examining on a population basis the changes in the availability of foods and/or nutrients for most of the 20th century. Sources of error inherent in use of these data as an index of per capita consumption were discussed in Chapter II-B-1. An additional consideration concerning use of the index of per capita disappearance involves its use in combination with databases on specific diseases. The national disappearance data refer to the total population and cannot provide information on subsets of the population; however, data on disease often refer to specific subsets of individuals at high risk for a particular disease. In such cases, information merged from the two data sets to examine associations between diet and disease over time should be interpreted with caution.

Dietary intake methodology used to collect data in cohort studies or prevalence studies may change with time, often in subtle ways that affect the comparability of the data collected at different times. In general, changes in wording of questions or interview procedures can introduce systematic bias and increases in within-person variation. The data of Folsom et al. (1987) illustrate some effects of changes in dietary intake methodology on comparability of survey data. In their study, nutrient intake was compared in the Twin Cities Metropolitan Area in 1973-1974 and 1980-1982. Though a 24-hour dietary recall was taken at both times, the instruments did not produce exactly comparable data. First, differences between the food composition tables used at each time led to differences in the results obtained. This problem was reconciled by analyzing both sets of data using one food composition table. Second, in the 1973-1974 survey only, use of alcoholic beverages was specifically probed. As might be expected, the recall of age-adjusted daily alcohol intake was greater in the earlier survey.

For comparison of totally different methods such as quantitative daily consumption versus semiquantitative food frequency, results of comparison studies should be available to show the differences in outcome resulting when the two methods are applied serially to the same people or cases. An example of such a study comparing results of food records
and semiquantitative food frequencies is that of Willett et al. (1985). Similar studies of the variation in a fixed set of food records resulting from use of different nutrient calculation systems are those of Jacobs et al. (1985), Polacchi (1985), and Taylor et al. (1985).

Use of different versions of tables of food composition may contribute to apparent differences in dietary intake data collected at different times. However, foods are continually added to and removed from the marketplace, composition of some foods has changed over time, and values for more nutrients have been incorporated into newer versions of food composition tables. In addition, coding rules (i.e., for the amount of fat lost in frying hamburger) may be changed. All of these changes may introduce artifacts to time trends. Biases will also be introduced if imputation procedures are not consistent for all time points.

It is appropriate to analyze dietary data according to the nutrient values that best characterize the foods when they were eaten, recognizing that linking date of food table entry to date of food consumption is logistically difficult. Knowledge of specific aspects of the food composition tables and coding practices is essential to avoid inaccurate comparisons resulting from changes in the databases over time. In the case of actual change in food composition, the Panel considers it preferable to use databases matched in time with each observation point; in the case of differences attributed to changes in analytical procedures, consistent use of one analytic value over all time points is considered more appropriate. At present there are no retrospective guidelines that permit users to make these types of adjustments.

2. Statistical problems

The basic statistical methodology for evaluation of time trends involves the two sample t-test (time one vs. time two) and its multiple group, regression, time series, and nonparametric analogues. These procedures involve correcting variance estimates for any correlations that might exist because of repeated measures on the same individuals. Correction for differences in baseline data may be appropriate in the comparison of nonrandomized groups in particular. In this case a "net difference" (change in one group minus change in the other) may be the desirable outcome measure that can be analyzed by t-test methodology. For classification problems (i.e., estimates of prevalence of nutrient intake level above or below a cutoff point), the two sample z-test for proportions provides the analogous statistical procedure.
Data interpretation is influenced by factors that may be considered as confounders or as explanatory variables. For example, changes in age composition of the population may be of no interest as a source of time trends in dietary intake. If so, results for each time period should be standardized to a common age distribution by direct standardization, indirect standardization, regression, or other method. However, age may be of interest as an explanatory factor in understanding time trends in dietary data. An increased proportion of pediatric or geriatric age groups may explain increased usage of certain food products. As another example, health education may change behavioral norms. This is a confounding factor in studying the relationship of a nutrient on a physiologic measure such as blood pressure but an explanatory factor in understanding why diets change.

High levels of intraindividual variation in dietary data may affect analysis of time trends. The power to detect changes in mean level of intake of a nutrient is reduced if the nutrient is measured imprecisely. Within-person variation itself does not bias the estimate of time trend; however, bias may be introduced in classification problems. Classification of the prevalence of nutrient level above a cutoff point is affected by sensitivity and specificity of analytical methods. Changes across time in these parameters may actually reverse the direction of a time trend (Quade et al., 1980; Whaley et al., 1980). Thus, an increase in sensitivity or a decrease in specificity might identify more "cases" at time two than at time one although there are actually fewer "cases" at time 2. This phenomenon would mask a real decrease in number of cases with time. Conversely, when a real increase in number of cases is present between the two times, an increase in specificity or a decrease in sensitivity may mask or even reverse the true time trend. Classification of prevalence is more precise (i.e., sensitivity and specificity both increase) if intraindividual variation is reduced. Statistical considerations in selection of cutoff points for determination of prevalence estimates are discussed by Brownie and Habicht (1984). Constancy in methods even when the initial methods are in some respect "wrong", may be preferable to changing to "better" methods because of the possibility of masking a real change.

Another area of concern for time trend analysis is survey variance. Specifically, variances are usually lower within a survey than between surveys when the individual is not the unit of analysis. This phenomenon has been observed in many settings where cluster sampling is used (Donner, 1984; Jacobs et al., 1986; Verma et al., 1980; Williams et al., 1981). A consequence of this large amount of variation is that an annual change in survey variables (including dietary variables) is often surprisingly large. However, this large change may be
reversed in the following year. For this reason, time trends consistent over several time periods should be observed before the time trend is accepted as real.

Analytical methods should generate variance estimates appropriate to the sampling method, or equivalently, to the unit of analysis. Single case studies, even when the single case is measurement of change in large units such as worksites, schools, or whole communities, (but with no control and no replication) should be interpreted with caution.

G. ERROR TERMS IN DIETARY DATA AND THEIR IMPACT IN ANALYSIS

Much attention has been directed toward the day-to-day variation of intake within an individual. In dietary data, this is a very large component of the observed variability of population intake and it is a component to be taken into account by methodological choice or analytical strategy at the stage of data analysis. Earlier in this chapter, the impact of "intra-individual variation" on various statistical analyses, and potential approaches to adjusting for this impact or adjusting observed distributions to reduce the contribution of this source of variation were discussed. Table 7, modified from the recent report of the Subcommittee on Criteria for Dietary Evaluation (National Research Council, 1986) illustrates the magnitude of this part of variance in relation to the variability that is of major analytical interest -- the variation between subjects. There are potentially major analytical implications when the intraindividual variation is as large as, or larger than, the interindividual variation, as is the case for almost all nutrients examined (Table 7). For a more extended discussion and for references to additional studies, see the report of the Subcommittee on Criteria for Dietary Evaluation (National Research Council, 1986).

Intraindividual variation is usually assumed to be a reflection of the real variability of an individual's intake from day to day. It is usually assumed that the variation is randomly distributed about a mean characteristic of the individual's usual or persisting intake. It is recognized that there may be a systematic effect of day of the week and, perhaps, of season. These must be taken into account in sampling days as well as individuals in population studies.

The studies reported in Table 7 and those reviewed by the Subcommittee on Criteria for Dietary Evaluation (National Research Council, 1986) strongly suggest that the variance ratio depends on at least the food component under study, the dietary methodology, the population group studied, and perhaps also the sex of the subjects. Because intraindividual variation
Table 7. Estimated Ratio of Intrinidividual:Interindivdual Variances Based on Reported Dietary Studies\(^1\) (National Research Council, 1986).

<table>
<thead>
<tr>
<th>Nutrients</th>
<th>24-Hr Recall/ Young Adults(^2)</th>
<th>3-Day Record/ Older Adults(^3)</th>
<th>1-Day Recall/ Year 1 (^4)</th>
<th>Year 2 (^4)</th>
<th>7-Day Record/ Adults(^5)</th>
<th>24-Hr Recall/ Pregnancy(^6)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>MALES</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Energy</td>
<td>1.1</td>
<td>1.0</td>
<td></td>
<td></td>
<td></td>
<td>0.8</td>
</tr>
<tr>
<td>Protein</td>
<td>1.5</td>
<td>1.2</td>
<td></td>
<td></td>
<td></td>
<td>1.4</td>
</tr>
<tr>
<td>Carbohydrate</td>
<td>1.6</td>
<td>2.1</td>
<td></td>
<td></td>
<td></td>
<td>0.6</td>
</tr>
<tr>
<td>Fat</td>
<td>1.2</td>
<td>1.2</td>
<td></td>
<td></td>
<td></td>
<td>1.3</td>
</tr>
<tr>
<td>Saturated Fatty Acids</td>
<td>1.1</td>
<td>2.2</td>
<td></td>
<td></td>
<td></td>
<td>1.4</td>
</tr>
<tr>
<td>Polyunsaturated</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Fatty Acids</td>
<td>2.8</td>
<td>3.5</td>
<td></td>
<td></td>
<td></td>
<td>1.9</td>
</tr>
<tr>
<td>Cholesterol</td>
<td>3.4</td>
<td>5.6</td>
<td></td>
<td></td>
<td></td>
<td>1.6</td>
</tr>
<tr>
<td>Vitamin A</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>1.6</td>
<td></td>
</tr>
<tr>
<td>Vitamin C</td>
<td>3.5</td>
<td>2.3</td>
<td></td>
<td></td>
<td></td>
<td>1.2</td>
</tr>
<tr>
<td>Thiamin</td>
<td>2.5</td>
<td>0.9</td>
<td></td>
<td></td>
<td></td>
<td>1.2</td>
</tr>
<tr>
<td>Riboflavin</td>
<td>2.4</td>
<td>0.9</td>
<td></td>
<td></td>
<td></td>
<td>1.2</td>
</tr>
<tr>
<td>Niacin Equivalent</td>
<td>1.6</td>
<td>2.2</td>
<td></td>
<td></td>
<td></td>
<td>1.1</td>
</tr>
<tr>
<td>Calcium</td>
<td>2.2</td>
<td>1.1</td>
<td></td>
<td></td>
<td></td>
<td>1.1</td>
</tr>
<tr>
<td>Iron</td>
<td>1.7</td>
<td>1.8</td>
<td></td>
<td></td>
<td></td>
<td>1.1</td>
</tr>
<tr>
<td><strong>FEMALES</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Energy</td>
<td>1.4</td>
<td>0.8</td>
<td>1.6</td>
<td>1.6</td>
<td></td>
<td>1.1</td>
</tr>
<tr>
<td>Protein</td>
<td>1.5</td>
<td>1.3</td>
<td>2.1</td>
<td>2.1</td>
<td></td>
<td>1.4</td>
</tr>
<tr>
<td>Carbohydrate</td>
<td>1.4</td>
<td>1.2</td>
<td></td>
<td></td>
<td></td>
<td>1.2</td>
</tr>
<tr>
<td>Fat</td>
<td>1.6</td>
<td>0.9</td>
<td></td>
<td></td>
<td></td>
<td>1.2</td>
</tr>
<tr>
<td>Saturated Fatty Acids</td>
<td>1.4</td>
<td>1.7</td>
<td></td>
<td></td>
<td></td>
<td>1.2</td>
</tr>
<tr>
<td>Polyunsaturated</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Fatty Acids</td>
<td>4.0</td>
<td>2.2</td>
<td></td>
<td></td>
<td></td>
<td>1.2</td>
</tr>
<tr>
<td>Cholesterol</td>
<td>4.3</td>
<td>4.2</td>
<td></td>
<td></td>
<td></td>
<td>1.1</td>
</tr>
<tr>
<td>Vitamin A</td>
<td>24.3</td>
<td>2.5</td>
<td>7.7</td>
<td>10.9</td>
<td></td>
<td>1.2</td>
</tr>
<tr>
<td>Vitamin C</td>
<td>2.0</td>
<td>2.8</td>
<td>2.3</td>
<td>2.5</td>
<td></td>
<td>3.9</td>
</tr>
<tr>
<td>Thiamin</td>
<td>4.4</td>
<td>1.6</td>
<td>3.3</td>
<td>3.9</td>
<td></td>
<td>3.3</td>
</tr>
<tr>
<td>Riboflavin</td>
<td>2.2</td>
<td>1.8</td>
<td>3.0</td>
<td>3.3</td>
<td></td>
<td>3.3</td>
</tr>
<tr>
<td>Niacin Equivalent</td>
<td>4.0</td>
<td>2.5</td>
<td></td>
<td></td>
<td></td>
<td>3.3</td>
</tr>
<tr>
<td>Calcium</td>
<td>0.9</td>
<td>1.7</td>
<td>1.1</td>
<td>1.2</td>
<td></td>
<td>1.0</td>
</tr>
<tr>
<td>Iron</td>
<td>2.5</td>
<td>1.5</td>
<td>2.7</td>
<td>2.5</td>
<td></td>
<td>1.0</td>
</tr>
</tbody>
</table>

\(^1\) Some estimates have been recalculated from original studies while other data were reported in form shown. All ratios represent the variances that would be seen with 1 day information. With replicated days, the ratios would decrease by the square root of number of days. The original papers may contain additional estimates. Only ratios available from two or more studies are included.

\(^2\) Beaton et al. (1979, 1983).

\(^3\) Hunt et al. (1983).

\(^4\) Sempous et al. (1985).

\(^5\) McGee et al. (1982).

\(^6\) Rush and Kristal (1982).

\(^7\) All of the variance attributed to intraindividual.
encompasses many components and therefore, may vary with specific methodology, with age or sex group, with income level or cultural practices relating to diversity of food selection, it is generally desirable to estimate the partitioning of variance within a particular study rather than inferring it from other studies (National Research Council, 1986). The Expert Panel noted also that day-to-day variation is not likely to be constant across individuals in a population. Adjustment procedures rely on estimates of the average variability. The limitations of this assumption, i.e., using a mean value, have not been developed although this would be desirable for the future. In the interim, the Panel notes that design of replicate observations should recognize this possible source of variation.

The suggestion that the dietary methodology (over and above the effect of number of days of data collected) influences the variance ratio leads to the awareness that the intraindividual variation includes not only the true biological variation in food and nutrient intake but also a number of other sources of variation which may differ with the methodology and food composition database adopted. These are discussed briefly in the paragraphs that follow.

1. Variance attributable to irregularity of consumption

Table 7 presents data for computed nutrient intakes. A general inference that may be drawn from the table is that an event that occurs infrequently or irregularly (intake of foods rich in certain nutrients such as vitamin A or polyunsaturated fatty acids) increases the day-to-day variance and the ratio shown in the table. Sempos et al. (1986) have recently reported that day-to-day variance is less for food groups consumed regularly than for food groups consumed irregularly.

Analyses of day-to-day variation in intake are not available for some of the other food components, additives, and contaminants in which FDA is interested. The operational consideration for analysis of such data is whether intake of the component of interest is likely to be reasonably consistent from day to day or to vary widely. The variance ratio is unlikely to be much lower than observed for energy (about 1); for many food components, additives, or contaminants, it could be appreciably higher. This will condition the constraints applicable to analysis and interpretation (see Chapter IV-A-D).
2. Variance arising in the estimation of food intake
   a. Under- and over-reporting of intake

In the absence of a reference "standard" method, it is difficult to assess the nature or magnitude of the effect of under- or overreporting of intake in a particular study. Errors in the estimation of true food intake on a given day can arise from several sources. The subject may forget to report intake, may consciously report erroneous intakes, or may incorrectly identify the actual food item ingested. The interviewer (or self-administered instrument) may probe for information to different degrees. Finally, in the coding of the original intake record, clerical errors in the identification of either the nature or quantity of food ingested may be introduced. Some of these error sources can be minimized by careful interviewer standardization and by close quality control measures. Studies of Beaton et al. (1979), Folsom et al. (1987), and Jacobs et al. (1985) illustrate differences in contributions to variance by these sources of error.

Given that it is possible, if not probable, that all of these contribute to observed variance, it is germane to consider the nature of their contribution.

b. Subject reporting error

Reporting errors can be random across days within an individual. In this case, the effect will be indistinguishable from what has been termed intraindividual variation and the analytical import will be the same as described earlier in this chapter. Reporting error can be systematically associated with other traits of individuals. For example, it has been suggested that overweight individuals tend to underreport intake and underweight individuals tend to overreport intake. Similarly, memory problems in elderly individuals may affect the ability to recall more distant intakes. In a total population study, these effects may remain random across individuals and will influence the apparent interindividual variation without biasing the population mean. However, if the population is stratified there may be a bias in the subgroup means. For the example suggested above, stratification of the population according to relative weight might result in systematic errors or biases within the strata.

c. Interviewer recording error

Similarly, interviewer effects can be random across days and subjects for an interviewer, can be systematic for a particular interviewer, or might represent an interaction
between particular interviewers and particular types of individuals. Depending upon the logistical strategy of the study, even an interviewer bias can become a random phenomenon if the assignment of interviewers-subjects-days is random. The greatest potential for systematic effects arises in studies in which different interviewers are involved, in which standardization of interview techniques and quality control are low, and in which subjects see the same interviewer on all occasions. A high degree of standardization and quality control may eliminate interinterviewer differences but at the same time may introduce a systematic bias across all data. In the case of self-administered instruments, this type of systematic effect, rather than "interviewer differences," is more likely.

d. Data coding error and coding rules

The coding of recorded food intake for entry into computer systems presents a range of possible types of errors that parallel the consideration of interviewer effects. The coder may be systematic or random in clerical coding errors. The effect on the coded data may differ depending upon whether all data for a particular subject are coded by the same individual. Another problem is presented by the operating "coding rules" established to deal with uncertainties in the recorded information (typically to address incomplete or ambiguous descriptors of the foods). These rules are necessary but they can introduce either a bias or a random error in the coded records. The nature and magnitude of the effect undoubtedly vary with the care with which coding rules have been established and with the quality control measures in place. The final effect can be quite small or appreciable (Beaton et al., 1979; Jacobs et al., 1985; Milner et al., 1983).

3. Variance arising from food composition data and computation

a. Missing data and imputation

The most serious impact of errors arising in the food composition databases is that of missing data. If this is treated in computation as representing zero, there will be an underestimation bias. If imputations are used to replace missing data, either random errors or bias can be introduced. Although imputations have clear impact on the estimation of a particular individual's intake, it has been demonstrated (National Research Council, 1986) that if the imputations represent random errors, they will not have a serious effect on univariate (distributional) analyses (see Chapter III-E).
b. Formulated foods and recipes

A special type of imputation is that associated with formulated foods or recipes for which composition is calculated from data on ingredients rather than from chemical analysis. The issues here are those of standardization of recipes (failing to trap the true variation; see food composition variability below); the manner in which cooking or processing changes have been taken into account; and, in processed foods, the variations in formulations that may be introduced in response to market price fluctuations. Two examples of changes in food composition resulting from fluctuating prices are the use of different vegetable oils in margarines and shortenings and the use of sugar or high fructose corn sweeteners as sweetening agents. Unless a food composition database is updated at frequent intervals and the food intake data are analyzed on a date-specific basis, there is no feasible way of incorporating these fluctuations in the computation of, for example, fatty acid or sucrose intake. Instead, the usual approach is to impute an "average" composition. In the Lipid Research Clinics' study (U.S. Department of Health and Human Services, 1982), the food composition database identified sucrose when it was a known component of a food, but classified it as "other carbohydrate" when it was not known whether sucrose or corn sweetener would be used at any point in time. While scientifically accurate in one sense, the effect is an underestimation of true sucrose intake.

c. Specificity of identification

Closely related to the composition of formulated foods is the specificity of identification of food items actually consumed. This can extend to the question of brand identification. Aggregation of information (use of one code and one composition estimate to represent an increasing range of items) eliminates variation from the computed data set. In effect, for some diets, intake of a component will be underestimated while for other diets it will be overestimated. Conversely, if the specification of food items is more precise than specification of items in the food composition database, the problems of missing values and imputations emerge. In many situations, precise identification is not necessary -- particularly if it can be assumed that the variation from the actual composition is random across items and brands. A potential problem arises in brand allegiance (e.g., in breakfast cereals) to products for which fortification or other compositional attributes show marked differences between brands (National Research Council, 1986). This is potentially a major issue for studies of intake of food additives or food contaminants that might be expected to be present in very different levels depending upon the manufacturer or perhaps the growing region. With respect to food additives or contaminants, the problem may be so great as to preclude the use of most food intake databases for such computation.
Regional differences for some food components may be important. These could arise from differences in production conditions (e.g., trace elements and soil composition, fatty acid composition of animals and feeding conditions) or from market differences (e.g., levels of added salt or color additives in different market areas, fatty acid composition of margarines in different regions) or, in the case of composite food coding, different typical mixes of food items/brand items in different regions.

d. Abridged food composition tables

The various abridged food composition tables that have appeared for use in conjunction with simplified dietary intake methods represent another variant on this issue. If the tables are well designed so that the aggregate composition datum is a good estimation of the average composition of the range of foods actually consumed, there will be no group mean bias associated with the use of abridged tables. However, of necessity, there will be some loss of variance in the computed data set. The extent of this effect undoubtedly depends upon the particular coding system. Grouping foods that are uncommonly eaten or have very similar compositions would be expected to have minimal final effect. Conversely, grouping foods that are commonly eaten and have quite diverse compositions for one or more single nutrients could have major impact for those nutrients.

A further variant on this same issue is the use of abridged food composition tables in conjunction with food frequency questionnaires. A frequency questionnaire can seldom capture either the complete intake of food or the full range of individual items consumed. It forfeits precision of identification and therefore, fails to capture part of the real variance in favor of avoiding the issue of true intraindividual variation in food and nutrient intake across days. Clearly, there is a trade-off of error terms that has both advantages and disadvantages depending on proposed analyses (see Chapter IV-A).

e. Biological variability in foods

Food composition is not fixed, even within a well specified food item. There is biological variability of the composition of individual items of the same class of food. In an examination of this phenomenon, the Subcommittee on Criteria for Dietary Evaluation (National Research Council, 1986) pointed out that the current USDA food composition data could be interpreted to suggest that the coefficient of variation of the composition of individual foods might be in the range of 10 to 45%. However, in a 1-day diet composed of perhaps 20 items, the relative error decreases, perhaps to a coefficient of variation of about 5 to 15%. This coefficient of variation was
estimated under the assumption that the variance is random across items of food and, therefore, that errors in individual items tend to cancel each other. Based on this logic, the variance across individuals would also be random. The net effect, would be that much of the unmeasured variance is really within the intraindividual component and might be reduced by averaging intake across days or by adjustment of distributions as discussed earlier in this chapter. However, a part of the unmeasured variance remains as interindividual variation because all individuals do not consume the same diet. A residual error appears, for example, as a small bias in estimated prevalence of inadequate intakes even after distributional adjustments (National Research Council, 1986).

f. Methodologic error

There is always a possibility of a true methodologic error of a systematic nature that will result in a bias toward under- or overestimation in data collection procedures or in food composition tables. Such changes should be documented in order to avoid mistakes in comparison of data among different studies or in estimation of time trends. As an example of such documentation, differences in food coding procedures and nutrient databases between the NFCS 1977-78 and the NFCS Continuing Survey of Food Intakes by Individuals were specified by the Human Nutrition Information Service (1985). Additionally, an example of a revision in food composition tables in recent years has been the lowering of the estimate of iron content of meats (Exler, 1983).

g. Derived databases of food composition

Many of the different databases of food composition used for estimation of nutrient intake originally drew on the same data but, through a series of specialized derivations, they have evolved separately. Updates in the original reference database may or may not have been transferred to the derived databases. Although data in some of the more widely recognized databases are maintained carefully, many databases in use within institutions are not regularly updated. Therefore, it is important to identify the database used in every study and to be able to document its origins and probable reliability. Comparative studies using reference 1-day intakes are available as part of the comparative validation process (Hoover et al., 1985).
4. Implications for analysis of various types of variances

The present section has examined a number of potential sources of "error" (here used in the statistical sense of a variance or deviation affecting the estimate of the true value that one wishes to estimate) in dietary data and the derived estimates of intakes of food components. Various potential or real sources of error have been identified. It is suggested that these could be translated into characteristic types of error that might or might not persist as error terms of significance in the final data analyses, depending upon how the error arises.

For purpose of discussion, the various error terms are categorized into three major groupings:

- errors that are random across days and across individuals;
- errors that may be random within the population as a whole but may associate in a systematic manner with certain individuals and, therefore, could be biases in certain subgroups of the population; and,
- systematic errors representing under- or over- estimations for all individuals and, therefore, representing a bias in the data set.

These are not the only error classifications that have relevance but they are useful in screening the potential impact for certain types of analyses. The underlying statistical principle is that random errors will tend to cancel out when mean values are derived whether these are means for a population of people or means of a population of days for individual subjects. Systematic errors will persist. These types of errors result in very different impacts on analyses.

For estimation of the intake of a particular individual on a particular day, all of the error sources, except the true day-to-day variation in intake, are real and applicable. Major error almost certainly exists in the estimated food component intake of a given individual on a given day. However, when estimates are made for analytical purposes of the "usual" intake of a particular individual or the distribution of usual intakes for a population group, the error terms may or may not be limiting for analysis. Table 8 summarizes the predicted impact of these errors on statistical analysis.

The present literature does not provide a systematic analysis of the types and magnitudes of error terms associated with particular dietary intake methods and food composition.
Table 8. Predicted Impact of Various Types of Error on Analysis*.

<table>
<thead>
<tr>
<th>Type of Error Term</th>
<th>Type of Analysis Planned (Based on &quot;Usual Intake&quot;)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td><strong>Univariate</strong></td>
</tr>
<tr>
<td>True Random</td>
<td>Can be greatly reduced by the adjustment of</td>
</tr>
<tr>
<td></td>
<td>distribution or by averaging intake across</td>
</tr>
<tr>
<td></td>
<td>multiple days. There will be some residual bias in prevalence estimates but not group means.</td>
</tr>
<tr>
<td>Random In Strata</td>
<td>Potential as above for total population studies.</td>
</tr>
<tr>
<td>Bias across Strata</td>
<td>Possibility of serious bias in group mean for substrata of the population.</td>
</tr>
<tr>
<td>Bias</td>
<td>There will be bias in the group mean and in the distributional estimates.</td>
</tr>
<tr>
<td>All</td>
<td>The confidence intervals for all estimates would be widened and statistical power would be diminished.</td>
</tr>
</tbody>
</table>

* For further details of the basis of these inferences, see Chapter IV A-D and the report of the Subcommittee on Criteria for Dietary Evaluation (National Research Council, 1986).
tables. Much of the past literature has focused upon comparison of methods to detect bias (i.e., validation of the group mean). When the distributions have been compared, the authors often have not taken due account of the effect of such obvious variables as the number of days of data collected. This situation is changing (Morgan et al., 1987a,b) and it is expected that future literature will provide more comprehensive analyses of error terms in dietary data.

Until that time, the method employed in a given study must be reviewed and informed judgments rendered about probable effect of methods on error terms in the data set and, therefore, on analyses of the data. The interpretation drawn from those analyses must be tempered by these considerations. There are important interactions between methodologic choice and analytical strategies: some methods yield data that are fully satisfactory for certain analyses yet totally unsatisfactory or seriously limiting for other analyses. The reverse is obviously equally true: certain analytical strategies are eminently suitable for data collected by one method and seriously misleading for data collected by another method. Matching data analysis strategies and food intake methods must be viewed as an important objective of the art and science of data analysis.
V. BIOLOGICAL INTERPRETATION OF ANALYSES OF DIETARY DATA

The foregoing discussions have focused upon limitations and/or potential uses of dietary data collected by various methods. In essence, a strategy has emerged by which the nature of the data and the nature of the statistical analyses can be paired in terms of acceptable and questionable matches. The consideration of the biological as contrasted to statistical interpretation of data and analysis has not been discussed in detail to this point. If the preceding chapters imply that the result of an analysis is likely to have a statistical bias, that bias carries forward as an error in biological interpretation. In addition, if the consideration of the quality of the dietary data suggests serious problems, these also carry forward. However, there are additional considerations of a purely biological nature. These are outlined below, together with suggestions about how they might be minimized.

A. BIOLOGICAL vs. STATISTICAL INTERPRETATIONS

Statistical error terms may be reduced by reversal of the dependent and independent variables in bivariate analysis. The case of simple regression, in which variable $Y$ is a physiological measure (such as serum level of a nutrient) with small variance relative to variable $X$, a measure of dietary intake, provides an illustration. If the variables are interchanged so that $X$ becomes $Y$ and $Y$ becomes $X$, then the measurement errors in diet are captured by the error term, the physiological measure is the measurement with lower variance, and the statistical problems are lessened. When variables are reversed for statistical analysis, biological interpretations must be made with great care.

Because of multicollinearity, some of the food components selected as "primary factors" by bivariate analysis may turn out to be either nonsignificant or reversed in multivariate analysis so that inverse associations become direct. This makes multivariate regression models difficult to interpret and statistical approaches cannot alleviate the difficulty (Reed et al., 1985). Gordon et al. (1984) suggest that dietary data should be interpreted on the basis of biological significance.

Although statistical models include and examine the contributions of single nutrients, it is well known that interactions of nutrients or food components can affect physiological status. Mixtures of nutrients may be a more critical factor in determination of biological effects than single nutrients (Freudenheim et al., 1986; Reed et al., 1985).
In all data analyses, statistical manipulations focus on description, removal of bias, and evaluation of the likelihood that observed values or differences of interest occurred by chance. Whether these observed values or differences have biological or other substantive implications is a separate matter. Findings that are statistically significant may not have biological importance. In this case, the study has identified a nonrandom effect. Conversely, findings may have biological significance without having statistical significance. In this case, the power of the test was inadequate for examination of a biologically important question. Therefore, both statistical and biological inputs are required for analysis and interpretation of data.

B. THE USE OF FIXED CUTOFF POINTS IN ASSESSING "RISK" OR PREVALENCE OF INADEQUATE NUTRIENT INTAKE

The customary approach for estimating the prevalence of inadequate nutrient intake has been the comparison of the observed intake of an individual with a predetermined cutoff point [usually set as a function of the Recommended Dietary Allowance (RDA)] followed by classification of the individual as having an "adequate" or "inadequate" intake (Beaton, 1987). This approach was criticized on several grounds (National Research Council, 1986). Several points related to interpretation are pertinent.

First, there is not a clear biological logic to the use of a constant proportion of the RDA as a cutoff point. Among nutrients, the RDA has different meanings (having been derived with differing underlying assumptions). Appropriate interpretation requires recognition that there is a distribution of requirements. Further, there must be an explicit definition of a "requirement for what" - i.e., if there is to be an estimation of the prevalence of inadequate intakes, then there must be an answer to the question "inadequate for what?" (National Research Council, 1986).

Second, the statistical issue of false positives and false negatives exists in the application of cutoff points. Given that requirements differ between individuals, no single intake level (requirement level) defines adequate or inadequate intake for all individuals even within a seemingly homogeneous class (e.g., young adult men). In the application of cutoff values, even to perfect dietary data, there will be misclassification of individuals. An unbiased estimate of prevalence can be obtained if the cutoff value is adjusted in such a way that the false positives and false negatives balance; otherwise, the prevalence will be underestimated or overestimated. It has been demonstrated that the perfect adjustment cannot be made until the prevalence is known.
(Brownie and Habicht, 1984; Habicht et al., 1982). The correct position of the cutoff depends on the relative positions and shapes of the intake and requirement distributions.

Third, the estimate of intake of the particular individual is likely to be erroneous unless there are multiple observations with acceptably low error terms. The concept of classifying a particular individual has serious limitations. Some of these concerns would be overcome if the population intake distribution were adjusted to reduce or remove the impact of intra-individual variation and cutoffs were applied to the adjusted distribution. The estimation of population prevalence might be improved, but this would not address the problem of the error of classification of particular individuals given a substantial variation in individual requirements.

C. USE OF A "PROBABILITY APPROACH" TO ESTIMATE THE PREVALENCE OF INADEQUATE NUTRIENT INTAKE

Application of fixed cutoff points may yield a bias in the estimate of the prevalence of inadequate intakes and may result in biological misinterpretation. In many cases utilizing a fixed cutoff point will overestimate the true prevalence of inadequate intakes.

Rather than using fixed cutoff values to estimate the prevalence of inadequate intakes, the Subcommittee on Criteria for Dietary Evaluation proposed a "probability approach" (National Research Council, 1986). In this approach, the likelihood that any observed level of intake is or is not adequate to meet the needs of a randomly selected individual is based on the relative areas under the requirement distribution to the right and left of the intake level. The probability estimates so generated may then be applied to the adjusted distribution of intakes. Probability of inadequacy times frequency of occurrence, summed across intervals of intake in the data set, provides an estimate of the prevalence of inadequate intakes. The estimate does not identify which individual's intake is presumed to be adequate or inadequate. The approach avoids the theoretical issue of the use of cutoff points and offers an approach to a "minimally biased" estimate of the prevalence of inadequate intakes.

The Subcommittee on Criteria for Dietary Evaluation cited a number of precautions with regard to the use of this approach (National Research Council, 1986). Care must be taken to express both intake and requirement in a manner that avoids spurious correlation between intake and requirement, i.e., observed intakes and requirements are independent. For
example, thiamin requirement is related to energy intake and metabolism; indeed, the primary estimates of thiamin requirements are referenced to energy intake. Thiamin intake also tends to increase with the total amount of food (and, therefore, energy) consumed. Thus, there must be a correlation between thiamin intake per day and thiamin requirement per day. If both are expressed as thiamin/1000 kcal/day, the spurious correlation disappears and there is a dramatic effect on the estimated prevalence of inadequate intakes. A parallel situation would be an analysis that did not separate the population under study into reasonably homogeneous age and sex groups; again, both intake and requirement might be expected to change in a systematic manner across widely different ages (e.g., between children and adults), giving rise to spurious correlations.

This approach to the provision of a "biologically interpretable" analysis of the adequacy of reported nutrient intake honors both statistical and biological interpretations. Estimates of mean nutrient requirements and their distributions must be generated before the approach can be applied (National Research Council, 1986). The feasibility of determining the distribution of requirements for nutrients varies markedly.

For comparison of intakes across two groups in the population (a presumed "at risk" group and a "nonrisk" group), it is not necessary to estimate prevalence. Statistical power is likely to be reduced if this is done. In that situation, either comparison of group means or comparisons of proportions below a fixed cutoff point may suffice (Brownie and Habicht, 1984).

D. EXAMINING THE OCCURRENCE OF EXCESSIVE INTAKE AND POTENTIAL TOXICITY

Although most investigators have been aware of the problems in nutritional assessment, the parallel issues in toxicological assessment are now being recognized more widely. Toxic levels of compounds are usually defined as 100 times below the lowest no observed adverse effect level (NOAEL) derived from animal experiments. Such toxicity levels are also used as legal standards to regulate the quantities of these compounds contained in foods. Toxicity standards for nutrients must be defined differently from toxicity standards for other chemical substances. For a nutrient, a hundredfold difference may not exist between the lowest NOAEL and the mean requirement for that nutrient. These issues are reviewed in the reports of the National Research Council (1986) and Allison et al. (1980).
It is beyond the scope of the present report to address the definition of criteria of inadequate or excessive intakes and the interpretation of cutoff points. However, the Expert Panel notes that the conceptual and statistical issues addressed in the report of the Subcommittee on Criteria for Dietary Intake (National Research Council, 1986) are applicable when the goal is to estimate the prevalence of "excessive" rather than "inadequate" intakes. A probability approach may be considered if the objective is to estimate the prevalence of excessive intake rather than the proportion of individuals with intakes above a semiarbitrary level.

When the objective is to estimate the prevalence of intakes above a legally defined "appropriate level of intake", rather than to estimate the magnitude of a biological problem, the use of a cutoff point, set at the acceptable level, is appropriate. It remains necessary to adjust the observed distribution of intakes to estimate the distribution of usual persisting intakes, if a statistically correct answer is to be obtained.

In some cases, estimates of the proportion of individuals with excessive intakes on any single day (e.g., an acute toxin) may be needed. Here it would be inappropriate to adjust the intake distribution to estimate the distribution of usual intakes. Rather, intakes for single days would be examined and the appropriate biological interpretation applied. The "probability approach" remains applicable. The random sources of error discussed in Chapter IV would not be canceled out because no adjustment of intake distributions or pooling of intake estimates across days for the same individual has been undertaken in these adjustments. Significant misclassification of individuals would occur because of these random errors (as well as any bias present in the intake data or derived estimates of food component intake) and the misclassifications might or might not cancel out. In all likelihood, some degree of bias would exist in the final results.

It is important to note that, biologically, individuals have not been classified correctly and only the proportion of individuals has been estimated with minimal bias. Particular individuals with persisting excessive intakes cannot be identified; nor, by inference, can the attributes of individuals who will be expected to have persisting excessive intakes be determined. In addition, as this approach is applied to food components other than nutrients, the inadequacies of existing food composition databases and the imprecisions of product identification in the original food records may impose serious limitations to interpretation. Food frequency questionnaires, based on target foods or classes of
foods, may be employed with advantage to circumvent the problems of estimation of persisting intake. If such is the case, the data may not correctly reflect total intake. The expected error would be one of underestimation of total intake or exposure.

Within any of these approaches, the prevalence of intakes is estimated over an arbitrary cutoff point set far below the accepted toxic levels. It cannot be inferred that the estimated proportion of the population will exhibit toxic manifestations.
VI. SUMMARY OF GUIDELINES

The ad hoc Expert Panel considers that the text of Chapters I through V comprises guidelines for the interpretation of dietary intake data. A summary of these guidelines is presented in the Executive Summary (p.v-xiv).
VII. LITERATURE CITED


Rosner, B.; Willett, W.C. [1987]. Interval estimates for correlation coefficients corrected for within-person variation: implications for study design and hypothesis testing. Submitted for publication.


VIII. STUDY PARTICIPANTS

ad hoc EXPERT PANEL

GUIDELINES FOR USE OF DIETARY INTAKE DATA

George H. Beaton, Ph.D.
Professor
Department of Nutritional Sciences
Faculty of Medicine
University of Toronto
Toronto, Ontario
Canada M5S 1A8

Frances A. Larkin, Ph.D.
Professor
School of Public Health
University of Michigan
Ann Arbor, Michigan 48109

David R. Jacobs, Jr., Ph.D.
Associate Professor
Division of Epidemiology
School of Public Health
University of Minnesota
Minneapolis, Minnesota 55455

Kiang Liu, Ph.D.
Associate Professor
Department of Community Health and Preventive Medicine
Northwestern University Medical School
Chicago, Illinois 60611

Nancy E. Johnson, Ph.D
Professor and Chairperson
Department of Food Science and Human Nutrition
University of Hawaii at Manoa
Honolulu, Hawaii 96822

Karen J. Morgan, Ph.D.
Associate Professor
Department of Human Nutrition, Foods, and Food Systems Management
University of Missouri
Columbia, Missouri 65211

Kathryn M. Kolasa, Ph.D.
Professor and Section Head
Nutrition Education and Services
Department of Family Medicine
East Carolina University
Greenville, North Carolina 27834

Milton Z. Nichaman, M.D., D.Sc.
Professor
Department of Nutrition and Epidemiology
University of Texas School of Public Health
Houston, Texas 77025

Helen Smiciklas-Wright, Ph.D.
Professor
Department of Nutrition
Pennsylvania State University
University Park, Pennsylvania 16802
REVIEWING CONSULTANT
Tim Byers, M.D., M.P.H.
Department of Social and Preventive Medicine
State University of New York at Buffalo
Buffalo, New York 14214

CENTER FOR FOOD SAFETY AND APPLIED NUTRITION, FDA
200 C Street, S.W.
Washington, D.C. 20204

Christine J. Lewis, Ph.D., R.D. Nutritionist
Clinical Nutrition Assessment Section
Division of Nutrition

Margaret McDowell, M.P.H. Nutritionist
Clinical Nutrition Assessment Section
Division of Nutrition

Elizabeth A. Yetley, Ph.D. Chief
Clinical Nutrition Assessment Section
Division of Nutrition

LIFE SCIENCES RESEARCH OFFICE

Kenneth D. Fisher, Ph.D. Director

Richard W. Leukroth, Jr., M.S. Staff Scientist

Sue Ann Anderson, Ph.D. Senior Staff Scientist

Judith F. Miller Administrative Assistant

Susan M. Pilch, Ph.D. Senior Staff Scientist

Steven H. Simpson Technical Literature Assistant

J. Elaine Huey Technical Literature Specialist

John M. Talbot, M.D. Senior Medical Consultant

Beverly R. Lea Technical Services Manager

Martha A. Watt Project Secretary
APPENDIX A

Scope of Work

Guidelines for Use of Dietary Intake Data

The following Scope of Work was assigned by the Food and Drug Administration (FDA) at the initiation of the project. As explained in the Introduction, the ad hoc Expert Panel refined the Scope of Work to more closely meet the needs of the FDA.

OBJECTIVE: To define guidelines for the correct analysis and interpretation of dietary data collected from large field surveys such as the National Health and Nutrition Examination Surveys (NHANES) and the Nationwide Food Consumption Surveys (NFCS).

The components shall include, but are not limited to the following:

1. Review past uses of the dietary data from large field surveys to gain a general understanding of the breadth of uses of such data.

2. Identify statistical procedures and other approaches for determining the accuracy and validity of data obtained through dietary surveys.

3. Provide guidelines for the correct approaches to statistical analyses of the data including consideration of complex survey design, if appropriate.

4. Identify appropriate uses of dietary data generated by different dietary assessment methods used in large field surveys (i.e., 24-hour recall, food frequency, etc.). Include a discussion of the usefulness of dietary data generated by different methods for:
   a. assessing nutrient status of individuals or groups
   b. determining diet-health relationships
   c. predicting health status (i.e., correlations with clinical and biochemical measures).
5. Identify inappropriate uses of dietary data generated by different dietary assessment methods used in large field surveys and indicate why these are inappropriate.

6. Provide guidelines for interpretation and use of dietary status assessments when information on other sources of intake are limited or lacking, e.g., information on vitamin/mineral supplement use, drugs, seasoning, and drinking water. Include consideration of the usefulness of different methods in minimizing biases due to these factors.
APPENDIX B

Glossary of Statistical Terms

The following definitions were adapted from Bourke et al. (1985), Steel and Torrie (1960), and Young (1981), or were employed as descriptive meanings for use with dietary intake methodology by the Expert Panel within the course of this study.

Accuracy

Degree of closeness of a measure to a standard or true value.

Bias

Measurement error. Bias may result from a systematic error, which tends to make the measurement consistently above or below the true value, or from random error, which tends to make the measurement randomly above or below the true value.

Experimental error

A measure of the variation that exists on experimental units treated alike. Sources of variation are: 1) the inherent variability that exists in the experimental material or subjects and 2) the variation that results from any lack of uniformity in the conduct of the experiment.

Precision

Agreement between the numerical values of two or more measurements that have been made in an identical manner.

Reliability

The repeatability or true reproducibility or error variation in collecting and processing dietary data.

Validity

The degree to which the method is a true measure of what the investigator wishes to describe.