Laboratory Issues: Use of Nutritional Biomarkers

Heidi Michels Blanck,*2 Barbara A. Bowman,† Gerald R. Cooper,‡ Gary L. Myers‡ and Dayton T. Miller‡

*Division of Nutrition and Physical Activity, National Center for Chronic Disease Prevention and Health Promotion and Epidemic Intelligence Service, Division of Applied Public Health Training, Epidemiology Program Office, Centers for Disease Control and Prevention, †Division of Diabetes Translation, National Center for Chronic Disease Prevention and Health Promotion, Centers for Disease Control and Prevention and ‡Division of Environmental Health Laboratory Sciences, National Center for Environmental Health, Centers for Disease Control and Prevention, Atlanta, GA

ABSTRACT Biomarkers of nutritional status provide alternative measures of dietary intake. Like the error and variation associated with dietary intake measures, the magnitude and impact of both biological (preanalytical) and laboratory (analytical) variability need to be considered when one is using biomarkers. When choosing a biomarker, it is important to understand how it relates to nutritional intake and the specific time frame of exposure it reflects as well as how it is affected by sampling and laboratory procedures. Biological sources of variation that arise from genetic and disease states of an individual affect biomarkers, but they are also affected by nonbiological sources of variation arising from specimen collection and storage, seasonality, time of day, contamination, stability and laboratory quality assurance. When choosing a laboratory for biomarker assessment, researchers should try to make sure random and systematic error is minimized by inclusion of certain techniques such as blinding of laboratory staff to disease status and including external pooled standards to which laboratory staff are blinded. In addition analytic quality control should be ensured by use of internal standards or certified materials over the entire range of possible values to control method accuracy. One must consider the effect of random laboratory error on measurement precision and also understand the method’s limit of detection and the laboratory cutoffpoints. Choosing appropriate cutpoints and reducing error is extremely important in nutritional epidemiology where weak associations are frequent. As part of this review, serum lipids are included as an example of a biomarker whereby collaborative efforts have been put forth to both understand biological sources of variation and standardize laboratory results. J. Nutr. 133: 888S–894S, 2003.

KEY WORDS: • biomarkers • diet assessment • epidemiology • nutrition • methodology

Nutritional biomarkers can serve as measures of nutritional exposure, or to use the nomenclature of environmental biomarkers, internal dietary dose. It is in this latter context that nutritional biomarkers go beyond being indicators of dietary intake and aid in our understanding of causal mechanisms between diet and disease. In 1983 Solomons and Allen (1) described the fundamental role of biochemical parameters, which today we call nutritional biomarkers, in assessing nutritional status. Their classic paper focused on choosing appropriate biochemical parameters (direct measures versus functional assays) and confounding factors and assigning diagnostic cutoff values. Since 1983 there have been many technical advances in the area of biomarkers as well as breakthroughs in the areas of genetics and metabolism.

Advances in these fields have been critical to nutritional researchers because biomarker levels can vary with absorption, metabolism, genetics and disease status. In addition to these advances, there has also been much discussion about issues related to proper sample collection (biological medium, fasting, contamination, stability), transport and storage (stability, temperature, light, oxygen) and analytical laboratory technique (precision, trueness, detection limit, recovery, standardization and quality assurance). Some of these issues are also discussed in this series by Potischman (2). Chapters by Hunter (3), Bates et al. (4), Sauberlich (5) and Myers et al. (6) and comprehensive articles by van’t Veer (7) and Hermus et al. (8) provide further reading on laboratory issues associated with nutritional biomarkers, which is the focus of this article.

When choosing a biomarker, it is important for the researcher to understand how a nutritional biomarker relates to both dietary intake and the chronology of exposure. This includes discussion of whether the biomarker will be used to evaluate long-term nutritional status, recent dietary intake, effectiveness of dietary manipulation or the efficacy of an intervention. These issues have been reviewed elsewhere (7)
and are also addressed in this series (2). In addition, it is important to know about the various types of biochemical tests that may be useful for nutritional status assessment, e.g., direct or static measurements, metabolite measurements, functional tests, abnormal metabolites, products of the nutrient under study, load or saturation tests or other procedures such as using stable isotopes (5).

Separate from essential knowledge about intake, chronology and biochemical tests is an understanding of how nutritional biomarkers are affected by sampling and laboratory procedures. In this laboratory context, at least four methodologic considerations should be taken into account when choosing an appropriate nutritional biomarker: 1) validity (how well the biomarker is measured in relation to its true value); 2) precision (how repeatable is the measure); 3) sensitivity (how well does the biomarker identify individuals with the condition); and 4) specificity (how well does the biomarker identify individuals without the condition) (9). The first two aspects assess how well the measurement can be made and the latter two explain how well the result can be interpreted.

How well the measurement can be made is partially dependent on the error associated with specimen collection and the analytic measurement of the biomarker. The following sections address these issues.

Measurement error: definition

Possessing validity or trueness means that the biomarker measures the relevant exposure accurately. This measure of accuracy is also called measurement error; it is the difference between the true value of the biomarker and the measured biomarker. To test validity, a gold standard or reference method that provides a good measure of the true exposure is necessary. This gold standard should reflect the true value without (or with minimal) laboratory or other sources of error. For a validation study to be appropriate, both the measured biomarker and the reference method must be related to the relevant nutrient exposure (10). Validation of biochemical measures has been reviewed by Hunter et al. (3) and Van't Veer (7). In most circumstances, measurement error is classified as either biological (preanalytical) or laboratory (analytical) error. Briefly, preanalytical error usually includes both biological and sampling errors, whereas analytical error focuses on the laboratory and includes method, instrument, reagent and/or matrix effects.

Preanalytical measurement error: genetics, environment, behavior and health status

Common sources of within-subject preanalytical biologic variability may arise from an individual's genetics, environment, behavior and health status and are therefore intraindividual or within-subject sources of error. A good review of bodily processes that may affect the availability of nutrients is found in the work of Bates et al. (4). Sources of variation include genetic polymorphisms; environmental and behavioral sources of variability can be from an individual's age, sex, diet, adiposity, weight loss, medication use, smoking status, physical exercise and alcohol intake. An example of variability from one's health status could include the case where the concentration of a specific nutrient affects the levels of another nutrient of interest, i.e., nutrient interaction.

Health-status variability can also occur when one's hormonal status or disease status influences a nutritional biomarker. This latter issue makes it important for researchers to be able to discriminate between the impact of diet on disease and the influence of disease on nutritional biomarkers.

The effect of disease on biomarkers of micronutrient status has been reviewed by Thurnham (11). Examples in the literature include reduction of plasma retinol, α-tocopherol and ascorbate by trauma; reduction of retinol and vitamin E due to changes in their transport proteins or lipoproteins during infection or other disease; an increase in vitamin C levels related to the leukocytosis common in trauma and changes in measures of iron status including serum ferritin, transferrin receptor and retinol binding protein as part of the acute phase response of chronic disease.

The impact of disease on biomarkers of macronutrient status (e.g., protein, energy) has been reviewed by Shenkin (9). Because of the possible effects of disease on biomarker level, researchers should determine a priori whether biomarkers of interest should be measured concurrently with measures of infection or disease status. This could include measures of malnutrition or protein status (total body or specific acute phase proteins), presence of inflammation or infection (malaria smears, erythrocyte sedimentation rate, C-reactive protein, white blood cell count, transaminase levels) and presence of disease such as hypertension, diabetes, nephrosis or myocardial infarction. Having this information allows the researcher to better use and interpret the biomarker data.

Preanalytical measurement error: sampling error

In addition to these potential sources of error, measurement error due to within-subject preanalytical variability can be due to differences related to sampling time. This variability within subjects can vary from hour to hour or year to year. Short-term variability may be hour to hour or day to day and can be caused by hormonal changes (e.g, menstrual cycle phase), fasting status or diurnal variation. An example of diurnal variation occurs with fasting glucose. If the accepted diabetes diagnostic cutpoint of 7.0 mmol glucose/L serum (126 mg glucose/dL serum) is used with afternoon samples instead of blood samples obtained in the morning after an overnight fast, approximately half of all cases of diabetes will be missed (12). Medium-term variability can be month to month and be due to seasonal changes in the diet such as those seen with cholesterol (13). Long-term variability may be year to year and due to intentional dieting patterns or changes in health status.

Analytical measurement error

Common sources of analytical or laboratory variability arise from errors in specimen collection and storage, errors during specimen analysis and inherent matrix effects from reagents, instruments and interfering substances (14). In addition repeated freeze-thaw cycles of serum samples can lead to increased variability (15). In an article by White (16), a number of potential laboratory sources of variability are reviewed. These include errors or omissions in the collection or analytical protocol. This type of error can occur as a failure to comply with the proper protocol for specimen collection (e.g., using the appropriate type of anticoagulant or preservative container, allowing the desired whole blood to clot, allowing sufficient clotting for serum, centrifuging in a timely fashion, keeping the specimen away from light or oxygen or at correct temperature), transportation or storage issues (keeping the specimen frozen or away from light or oxygen). Every effort should be made to standardize the specimen-collection protocol; however in a study that covers a wide geographic area or a long period of time, it may not be possible to treat all samples exactly the same (i.e., some will travel further and some will be stored for longer periods of time).

Measurement errors due to variations in analysis of the biomarker may also result from variability in technique and motivation between laboratory technicians, use of different or contaminated reagents and failure to maintain standardization of the instrument during the course of data collection. Once standardization is accomplished however, it is important to maintain quality control, but it is not necessary to frequently
restandardize. See the work of Petersen et al. (17) for a discussion of proposed guidelines for internal quality control in the medical laboratory.

**Measurement error: assessment**

Precision or reproducibility implies reliability and expresses the variability of results obtained from a single sample measured many times (4). Most laboratory researchers express the random analytical error of measurement as the standard deviation determined by repeated measurements performed on the same subject sample. Usually a mean and standard deviation (SD) are calculated from several measurements on the sample and are then used to calculate the coefficient of variation (CV); \[ CV = \frac{SD}{\text{mean}} \times 100. \] The CV is ideally calculated for samples at the bottom, middle and top of the reference concentration range that was determined on healthy people. Laboratories can provide both within-run and between-run CV to the researchers. The within-run CV is determined by dividing the samples into two or more aliquots and analyzing them together. The between-run CV is obtained when aliquots are analyzed in different runs, which is usually on different days. Multiple publications have presented approaches for estimation of reference intervals from different numbers of specimens and numbers of analyses (18).

**Measurement error: minimization**

A major approach to minimize laboratory measurement errors involves blinding the laboratory analyst to the disease status of the specimens or other pertinent characteristics and eliminating systematic differences in the way specimens are handled. As reviewed by Hunter (3), ideally all specimens should be analyzed in a single run to reduce between-assay variation (laboratory drift). In addition, if possible, paired-case and control samples should be analyzed consecutively (with the within-pair order randomized) and with the pairs ordered randomly with regard to other variables of interest so that effects of order (within-run laboratory error) are not attributed to another variable. If all samples cannot be assessed in a single run, then cases and an appropriate number of controls should be analyzed together in the same batch to ensure the validity of the paired comparison.

For some nutrients, certified quality controls and reference samples are available to help ensure comparability of results obtained by different laboratories (Table 1). Some standards are certified by the National Institute of Standards and Technology and are made available as Standard Reference Materials. If suitable reference standards are not available, researchers should test a laboratory by having the staff measure duplicate samples with laboratory personnel blinded. Such samples should be included within the sample batches to monitor drift and reliability. Also it is essential for each laboratory to establish its own internal quality controls, for example, by including aliquots of known well-characterized deficient and normal controls for which mean and standard deviation values have been established. These aliquots can also be exchanged with other laboratories for comparisons of results. Routine inclusion of at least two samples from two quality control or reference samples in every batch of specimens helps to maintain precision and prevent or reduce between-assay variation, i.e., drift. In addition use of these reference materials over the entire range of possible values helps to assess the accuracy of the method. A list of available standards and secondary reference materials for lipids and lipoproteins has been published (6,19).

**Measurement error: effect in epidemiologic studies**

Measurement error can lead to bias in measuring the association between nutritional exposure and outcome. This bias is known as either information bias or misclassification bias. In epidemiologic studies, it is important to minimize misclassification bias from measurement error and to determine whether this bias is differential or nondifferential. The effects of measurement error in epidemiologic studies are discussed in recent reviews by Saracci (20) and White (16), in the text by Willett (21) as well as in this series (22).

Differential error occurs when the measurement error differs between those with and those without the outcome or disease of interest. The influence of this type of error on the measure of association can only be evaluated if information about the error is known for all groups. However, if the presence of disease is in fact the cause of changes in a biomarker, it may falsely lead one to conclude that there is an association between biomarker and disease. To address this problem, stratification by disease stage can be a way to determine whether disease-induced differential error exists (20).

Nondifferential error occurs when the error does not differ between the comparison groups. This type of error is generally tolerable but can cause the measure of association to be attenuated toward the null, which results in a lack of effect between the nutritional biomarker and the outcome of interest. This can occur in a study where the overall CV for the nutritional biomarker of interest is high. In general, as the analytical precision decreases and therefore the CV increases, the odds ratio is attenuated toward the null. This can lead to an incorrect interpretation, i.e., the lack of an association between biomarker and outcome. It has been suggested that for epidemiological studies ideally the CV should not be > 5% (3). This level of accuracy is very difficult to attain for many nutrients; often it is not possible to attain a CV < 10%. At levels > 10% there may be concern about the utility of the assay. In all cases, inclusion of the CV for the assay in the report of the study results is important. The nationally accepted CV cutpoint for total cholesterol is 3%; low-density lipoprotein (LDL) cholesterol, 4%; high-density lipoprotein (HDL) cholesterol, 4%; triglyceride, 5% (6).

### Table 1

<table>
<thead>
<tr>
<th>Examples of nutritional biomarker reference sample and certified quality control sources</th>
</tr>
</thead>
<tbody>
<tr>
<td>National Institute of Standards and Technology Standard Reference Materials (U.S.A.)</td>
</tr>
<tr>
<td>National Institute of Biological Standards and Controls (U.S.A.)</td>
</tr>
<tr>
<td>World Health Organization, Blood Safety and Clinical Technology</td>
</tr>
<tr>
<td>International Federation of Clinical Chemistry, Scientific Division</td>
</tr>
<tr>
<td>Centers for Disease Control and Prevention, Division of Laboratory Sciences</td>
</tr>
<tr>
<td>Northwest Lipid Research Laboratories, University of Washington, Seattle, WA</td>
</tr>
<tr>
<td>Solomons Park Research Laboratories, Kirkland, WA</td>
</tr>
<tr>
<td>Commercial companies (primary standards)</td>
</tr>
<tr>
<td>Proficiency testing programs</td>
</tr>
<tr>
<td>National reference material institutions</td>
</tr>
</tbody>
</table>

---

3 Abbreviations used: apo, apolipoprotein; CRMLN, Cholesterol Reference Method Laboratory Network; CV, coefficient of variation; HDL, high-density lipoprotein; LDL, low-density lipoprotein; NCEP, National Cholesterol Education Program; NHANES, National Health and Nutrition Examination Surveys.
Misclassification bias can also come from the use of inappropriate biomarker cutoff values. One must understand the method's limit of detection and laboratory cutoff points to maximize the biomarker's sensitivity and specificity. Overlap of persons with low or deficient nutritional biomarker values and those normal values can lead to misclassification of individuals thereby affecting the sensitivity and specificity of the biomarker. As reviewed by van den Berg (23), misclassification can result from confounding factors but can also be caused by both biological and laboratory variability; therefore minimization of sampling and laboratory errors are important in this aspect of analysis as well. To aid in the choice of a biologically relevant cutoff value, one can use serial or multiple measurements or a combination of nutritional exposure measurements, which is called a composite value. This latter choice, although affected by individual measurement error, may overcome some of the imperfection of nutritional biomarkers in epidemiology. As a final note, if the biological variability is such that nutrient levels show higher within-subject (intraindividual) variability as compared to between-subject (interindividual) variability, the researcher must be aware of the diminished power of this biomarker and consider increasing the sample size of the study or, as mentioned, consider using a combination of nutritional exposure measurements.

Separate from issues of measurement error, use of a percentile cutoff value as determined from a healthy population distribution may lack any real biological or functional basis. Similarly using percentile cutpoints from an even distribution of the survey's participant values may lead to null results. Therefore functional tests or the use of specific subpopulation data such as elderly, young, adolescent or pregnant women may be of value in determining the optimal cutoff value (23). The National Health and Nutrition Examination Surveys (NHANES) are an example of a reference data source that allows for determination of subpopulation reference ranges and cutoff values for a healthy U.S. population by sex, racial/ethnic group and age group.

It should also be noted that many epidemiologic analyses use continuous-exposure data. Appropriate regression diagnostic approaches should be used to check the assumption that use of a continuous dependent variable is warranted, and one also needs to determine whether a one-unit change in the dependent variable is a meaningful way to interpret the data. Graphing of the data should be used to ensure that a linear association is appropriate. It may be necessary to transform the nutritional values, i.e., log transformation $Y' = \log Y$, square transformation $Y' = Y^2$, to normalize the dependent variable, to stabilize the variance of the dependent variable or to linearize the regression model if the original data suggests that the relation is nonlinear.

Another problem inherent in nutritional biomarker research is the possibility that biologic processes such as genetic polymorphisms can alter biomarker values. This phenomenon is termed metabolic confounding by Saracci (20). If it is known that individuals may have different forms of an enzyme that can ultimately affect biomarker metabolism, then the enzyme activity should be considered as a potential confounder and controlled for. One way to control for this is by stratifying on the enzyme status. An example of a genetic polymorphism of this type is human apolipoprotein (apo)E, which has three common alleles (E2, E3 and E4). These alleles code for three isoforms that are associated with different levels of serum cholesterol (24). Similarly an inverse relationship has been found between apo(A) isoform size and plasma lipoprotein(A) levels (25).

Quality control and calibration

The aim of quality control is to ensure that the analytical values produced by a laboratory are sufficiently reliable for their intended purpose (26). A good quality control program monitors the following: 1) preanalytical variation; 2) clerical error—proper labeling and logging in of specimens as well as maintenance of appropriate records for all specimens for future reference; 3) technique—assurance that all laboratory analysts understand the principles that underlie a specific assay and that all personnel use the same technique and have ready access to a detailed and current method manual; 4) reagents and materials—proper labeling, evaluation for matrix effects in analytical system, confirmation of absence of interfering substances, substitution of new reagents does not change level of values of quality control materials; 5) calibration—using a purified primary standard or a commercially available serum secondary calibrator traceable to an accepted reference method or reference material, confirmation that each newly acquired calibrator possesses an accurate target value; 6) bench performance—using controls and standards for each assay, documentation of daily bench performance for detection of slight error over time, and 7) instrumentation—performance of periodic preventive maintenance for all instruments with appropriate maintenance documentation.

Some methodological problems associated with certain laboratory assays are listed in the appendices of the Dietary Reference Intake reports (27), where tables summarize available data on different methods: accuracy and precision, analytic sensitivity and specificity, interlaboratory agreement and how changes in methods have affected estimates over time. However, individual laboratories should be able to provide internal and external quality control data for biomarker assessment.

Gunter et al. (28) have published an article on the results on an international round-robin for serum folate and whole blood folate. This interlaboratory comparison study was conducted to assess differences among methods. Twenty research laboratories participated in a 3-d analysis of six serum and six whole blood pools. Overall mean, standard deviation and CV values derived from these results were compared within and across method types. Results reported for serum and whole blood folate resulted in overall CV of 27.6 and 35.7%, respectively, across pools and two- to ninefold differences in concentrations between methods with the greatest variation occurring at low folate concentrations. Although results for serum pools were less variable than those for whole blood pools, substantial intermethod variation occurred. These results emphasize the urgent need for developing and validating reference methods for biomarker measurements and for properly characterized reference materials. In addition these findings highlight that when one is evaluating study or clinical data, method-specific reference ranges (established with clinical confirmation of values for truly deficient individuals) must be used.

In addition to the interlaboratory comparison study performed on serum and whole blood folate, Pfeiffer et al. (29) compared plasma total homocysteine measurements in 14 laboratories. These 14 laboratories across the world used eight different analytical methods. The laboratories participated in a 2-d analysis of 46 plasma samples and plasma quality control pools. The mean among-laboratory and within-laboratory (among-run) CV were 9.3 and 5.6%, respectively, for plasma samples. Specific techniques had systemically higher or lower values when compared to gas chromatography/mass spectrometry. This analysis showed that among-laboratory variations within one homocysteine method could exceed among-method...
variations. Although some of the methods appeared to be interchangeable, these findings further draw attention to the need for improved analytical precision.

Quality control: long-term planning

For researchers involved in long-term prospective studies, consideration of long-term quality control is important. A review of long-term laboratory quality control planning has been done by researchers at the NHANES nutritional laboratory (30,31). As reviewed, long-term quality control not only encompasses methodologic accuracy and precision but also must ensure positive specimen identification (including adhesion of labels after liquid nitrogen or boiling-water bath immersion) and address time-associated trends such as specimen stability (including rate of deterioration over time and freeze-thaw effects), changes in analysts, reagents or instrumentation and seasonal/geographical variation. For example, at the NHANES lab, both bench-quality controls prepared at low, normal and elevated concentration levels are run two to four times in each analytical run, and blind-quality controls and low and moderately elevated normal levels are incorporated into every run.

To assist the reader, Tables 2A–C list considerations for choosing the appropriate nutritional biomarker, choosing a laboratory, and understanding the necessities of specimen collection and processing. For an example of a detailed review of choosing an appropriate assay and accompanying laboratory, see the article by Nexo et al. (32), which describes this type of evaluation for homocysteine.

Understanding sources of variation: results from the National Cholesterol Education Program

Serum lipids are an example of important biomarkers for which collaborative efforts have led to both understanding biological sources of variation and standardization of laboratory measurements (6). Cooper et al. (13) reviewed the magnitude and impact of the major biological and analytical sources of variation in serum lipid and lipoprotein levels on risk of coronary heart disease. Because of the concern regarding serum cholesterol and disease risk, concerted efforts were made to attain analytical accuracy and precision (e.g., CV < 3% by 1992) as well as to minimize biological and sampling sources of error. It was determined that behavioral sources of variation in serum lipids arise from the lifestyle of the individual. Lifestyle factors include diet, weight status, smoking, physical activity and alcohol consumption. In addition to behavioral sources of variation, there are also a number of clinical sources of variation in serum lipids. These sources include metabolic states, illness and secondary or other disease states such as myocardial infarction, stroke, hypertension, nephrosis, diabetes and

### TABLE 2A

<table>
<thead>
<tr>
<th>General considerations in choosing a nutritional biomarker for an epidemiologic research study</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Timing relative to dietary exposure: recent intake versus usual intake, acute versus chronic exposure.</td>
</tr>
<tr>
<td>2. Type of measurement: direct measure (static indicator) versus functional assay.</td>
</tr>
<tr>
<td>3. Would a dietary assessment method such as a food-frequency questionnaire or a 24-h recall provide adequate dietary information precluding the need for biomarker assessment?</td>
</tr>
<tr>
<td>4. Has within-person (intra-) and between-person (inter-) variance been documented for the biomarker measurement method of interest? If yes, is the between-person variance larger than the within-person variance? If no, it will be difficult to assess associations without an extremely large sample size.</td>
</tr>
</tbody>
</table>

a) Has the laboratory method been standardized against an accepted reference method?

b) Has the laboratory method been shown to be reliable/reproducible?

c) Have the specificity and sensitivity of the biomarker been established?

d) Do appropriate cutoff values or reference ranges exist for the biomarker within the population you are studying?

e) Can you assume there is a sufficient range of concentrations in the biomarker concentration within your sample to warrant its use?

f) Is it logistically possible to obtain the specimen as the protocol specifies or do collection and transportation conditions negate its use, i.e., store specimen away from light at 4°C?

### TABLE 2B

<table>
<thead>
<tr>
<th>Considerations in choosing a laboratory for assessment of a nutritional biomarker</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Does the laboratory have a written description of the entire preanalytical, analytical and reporting procedures for the analytes to be determined for the epidemiologic, clinical investigations and methodology research studies or for collaborative services in the clinical laboratory?</td>
</tr>
<tr>
<td>2. Does the claimed analytical performance (coefficient of variation and bias) of the laboratory meet the requirements necessary to accomplish the goals of the epidemiologic or clinical studies?</td>
</tr>
<tr>
<td>3. Confirm with acceptable quality control serum pools that the laboratory meets the claimed acceptable precision and bias requirements.</td>
</tr>
<tr>
<td>4. Confirm that the calibration process uses a purified primary standard of known purity or a commercially available serum calibrator that is traceable to an international or national accepted reference material. Inquire about recovery experiments and whether frequency of use of calibrator meets statistical requirements.</td>
</tr>
<tr>
<td>5. Confirm that the laboratory uses statistical quality control charts at the bench level to maintain accuracy among runs; alert of bias changes with new calibrators, new instruments or deteriorating reagents and provide analytical performance data over long-term periods.</td>
</tr>
<tr>
<td>6. Determine whether a clinical laboratory participates successfully in an effective proficiency program and Clinical Laboratory Improvements Amendments evaluations or whether an epidemiologic laboratory has proven its performance is traceable to highly accurately labeled concentration values of national or international reference materials.</td>
</tr>
<tr>
<td>7. Before analysis, ensure that the sample-collection process, sample preparation and storage conditions and patient medication or disease status do not cause matrix effects in the analytical instrument-reagent system, which can result in inaccurate results. Suspect matrix effects with frozen, lyophilized or deteriorating samples denatured by unsuitable storage, interfering substances and inappropriate handling techniques.</td>
</tr>
<tr>
<td>8. Confirm that highly accurate measuring instruments are being used for measurement of critical volumes, carry-over contamination does not occur in the spectrophotometer or absorption measuring instrument, the sample-handling procedure is reproducible and possible transcription errors are controlled.</td>
</tr>
<tr>
<td>9. What are the laboratory’s analytical reference ranges of the analytes? Is the laboratory report suitable for use in the study of the population or subgroup of individuals being assessed? What is the lab throughput and turnaround? Will the assays be completed in the time frame needed?</td>
</tr>
<tr>
<td>10. Is the cost of the assay appropriate and competitive for the demands on accuracy and precision, the necessity to use duplicate or more determinations, the cost of highly qualified personnel assigned to the project and the possibility that instrumentation and reagents are highly expensive?</td>
</tr>
<tr>
<td>11. Is the laboratory staff motivated toward high-quality analytical performance and does the staff receive close supervision?</td>
</tr>
</tbody>
</table>
infections. Sampling sources of variation in serum lipids also include fasting status, posture and choice of serum versus plasma.

The accuracy of the determination of risk for cardiovascular disease of a cholesterol measurement in clinical trials and other epidemiology studies depends on the availability of an accurate reference method and cholesterol reference materials labeled with accurate target values (33). The Centers for Disease Control and Prevention (CDC)-modified Abell-Levy-Brodie-Kendall cholesterol method is accepted as the cholesterol reference method of the National Reference System for Cholesterol. This method is also the reference method for HDL and LDL cholesterol (34).

An investigation of the effect of systematic bias and random error, quality control and intraperson biological variation on the National Cholesterol Education Program (NCEP) clinical classifications for reported lipid measurements has been performed. This investigation found that the NCEP guidelines are adequate to ensure correct patient classification at least 90% of the time if the laboratories are meeting the NCEP guidelines for analytical precision and are using standard quality control procedures (35). This probability is ensured regardless of the size of the systematic bias of the laboratory or increased random analytic error. In collaboration with manufacturers of cholesterol diagnostic products, validity of trueness of serum lipid measurements in the clinical laboratory is being improved by manufacturers’ calibration of clinical laboratory equipment to the reference methods maintained at the CDC through activities by the National Cholesterol Reference Method Laboratory Network (CRMLN) (36). The CRMLN uses CDC serum lipid reference materials labeled by the CDC reference method to assist manufacturers of in vitro diagnostic lipid assays with the validation of their lipid and lipoprotein assays before the assays are distributed to the clinical laboratories. This is done so that clinical laboratories across the U.S. can be confident that they can accurately measure total, HDL and LDL cholesterol.

**SUMMARY**

Measurement error is often classified as either biological (preanalytical) or laboratory (analytical) error. Preanalytical error usually includes both biological and sampling errors, whereas analytical error focuses on the laboratory and includes method, instrument, reagent and/or matrix effects. Common sources of laboratory variability arise from errors in specimen collection and storage, errors during specimen analysis and from differences in reagents, instruments and interfering substances. Two major approaches to reducing laboratory measurement error are to blind the laboratory analyst to the case-control status of specimens and to eliminate systematic differences in the way case and control specimens are handled. Reducing measurement error (either preanalytical or analytical) and choosing appropriate cutpoints is extremely important in nutritional epidemiology, where weak associations are frequent.

To correctly interpret the association of nutritional exposure and outcome, it is also important to anticipate the nature of the errors. It is the investigator’s responsibility to choose a valid biomarker and to interpret results with knowledge of the measured error. This error should be included in the reporting of the results to aid the readers’ interpretation of the findings and comparison with other studies.

Although the use of nutritional biomarkers in screening was not specifically addressed in this paper, similar issues and concerns must be heeded. In addition previous knowledge or estimation of sensitivity and specificity for the disease of interest is needed to ensure correct interpretation of the results. The interpretation is important for proper implementation of treatment and/or control measures.

Future studies on the impact of disease and genetics on nutritional biomarker levels will help researchers better interpret measured biomarkers and hence aid in the understanding of the causal role between diet, metabolism and genetics and disease. As assessment of nutritional biomarkers improves, they may be used to assess dietary exposure as well as provide important information on early biological effects and disease characteristics. Determining and quantitating the sources of biological and laboratory error for all nutritional biomarkers (such as has been done for lipids) will therefore be an important step in the area of nutritional epidemiology.

**LITERATURE CITED**


---

**TABLE 2C**

**Considerations in nutritional biomarker specimen collection and processing**

1. What is the correct specimen type to collect (e.g., plasma, serum, urine, bloodspots)?
2. What is the appropriate method of collection (e.g., capillary sample, venipuncture, midstream urine, etc.)?
3. What is the appropriate vial? Vials must hold appropriate volume, be stable when frozen and hold their contents without leakage; label or markings must have long-term adhesion and have integrity at a range of temperatures.
4. Is a preservative or anticoagulant required?
5. What volume of sample is required?
6. Will dilution, preservation, aliquoting or hemolysate preparation of the sample be needed?
7. How large should the aliquots be to maximize the efficiency of use of the samples and to meet requirements for the laboratory assay?
8. Do the subjects need to fast or follow other instructions before sample collection?
9. Are samples sensitive to light, oxygen or temperature?
10. Do you have access to a local laboratory or can you transport collection materials to the field (i.e., centrifuge, refrigerator or freezer, wet or dry ice, phlebotomy supplies including gloves, alcohol or alcohol pads, gauze, bandages, Vacutainer tubes and accompanying needles, barrels and biohazard disposal containers)?
11. Will you be able to collect field controls?
12. Do you have enough supplies to split the samples to produce duplicates?
13. Is the stability of the sample known under field conditions? At what optimal temperature should the samples be transported (e.g., must the samples be shipped refrigerated within 24 h of collection or can they be frozen or refrigerated and then shipped or transported on wet or dry ice)?
14. At what optimal temperature should the samples be stored? What is the optimal size for storage aliquots to maximize utility and minimize loss of sample?