Guidelines for Data Acquisition and Data Quality Evaluation in Environmental Chemistry

ACS COMMITTEE ON ENVIRONMENTAL IMPROVEMENT

Daniel MacDougall, Chairman
General Manager
Dow Pharmaceuticals—Canada

Francis J. Amore
Senior Research Chemist
Henkel Corporation

Geraldine V. Cox
Technical Vice President
Chemical Manufacturers Association

Donald G. Crosby
Professor, Department of Environmental Toxicology
University of California

Francis L. Estes
Consultant for Environmental Sciences

David H. Freeman
Professor of Chemistry
Department of Chemistry
University of Maryland

William E. Gibbs
Vice President and Director of Research and Development
American Hoechst Corporation

Glen E. Gordon
Professor of Chemistry
Department of Chemistry
University of Maryland

Lawrence H. Keith
Manager
Analytical Chemistry Division
Radian Corporation

Joginder Lal
Manager, Exploratory Polymers
Goodyear Tire & Rubber Company

Ralph R. Langner
Director, Industrial Hygiene Laboratory
Dow Chemical U.S.A.

Nina I. McClelland
Vice-President, Technical Services
National Sanitation Foundation

Wendell F. Phillips
Senior Research Scientist
Campbell Soup Company

Robert B. Pojasek
Technical Director of Hazardous/Toxic Substances Management
Weston Inc.

Robert E. Sievers
Professor and Co-Chairman
Department of Chemistry
University of Colorado

Committee Staff

Robert G. Smerko
Director, Department of Public Affairs
American Chemical Society

David C. Wimert
Committee Liaison
Department of Public Affairs
American Chemical Society

SUBCOMMITTEE ON ENVIRONMENTAL ANALYTICAL CHEMISTRY

Warren B. Crummett, Chairman
Technical Manager, Analytical Laboratories
Michigan Division
Dow Chemical U.S.A.

Francis J. Amore
Senior Research Chemist
Henkel Corporation

David H. Freeman
Professor of Chemistry
Department of Chemistry
University of Maryland

Robert Libby
Associate Director
Corporate Technology Division
Proctor & Gamble
Miami Valley Laboratories

Herbert A. Laitinen
Graduate Research Professor
Department of Chemistry
University of Florida

Wendell F. Phillips
Senior Research Scientist
Campbell Soup Company

Michael M. Reddy
Research Scientist IV
Geological Survey
U.S. Department of Interior

John K. Taylor
Coordinator for Quality Assurance and Voluntary Standardization
Center for Analytical Chemistry
National Bureau of Standards

PREFACE

Sound environmental analytical chemical measurements are essential to provide the data needed to ensure the quality of the environment and the health of the public. In September 1978 the American Chemical Society's Committee on Environmental Improvement expressed concern that current practices used in environmental analytical chemistry varied so much that results were not useful for interlaboratory comparisons. Furthermore, the Committee questioned the validity of some analytical methods used to generate data; the analytical methods used often are not written in a form complete enough to allow others to understand how the analysis was conducted.

The Committee has directed its Subcommittee on Environmental Analytical Chemistry to develop a set of guidelines

0003-2700/80/0352-2242$01.00/0

that would help improve the overall quality of environmental analytical measurements. The following "Guidelines for Data Acquisition and Data Quality Evaluation in Environmental Chemistry" provide guidance in developing reliable chemical analysis of environmental samples and allow more confident interlaboratory use of the data.

It should be recognized that environmental analytical measurements are developed for a variety of purposes, such as, the determination of the environmental fate (transformation and transport) of a chemical and determination of the environmental concentration of a chemical, for use in environmental risk assessments or in some cases for regulatory purposes. This broad range of application of analytical data and the variety of decision-making processes may require differing ranges of analytical certainty. However, the accuracy must be adequate for each use, and this must be established in every case. The Committee believes that the guidelines are generally useful and should be broadly applied throughout the field of trace organic analysis and urges that they be used for regulated samples.

While we recognize that the field of analytical chemistry is advancing toward more precise and accurate methods of analysis, the equipment, personnel capabilities, and availability of new analytical techniques vary from laboratory to laboratory. The Committee hopes that these guidelines will provide the basis for a uniform approach throughout the environmental analytical field and will allow improved interlaboratory comparisons. It also recognizes that the guidelines will require updating as scientific knowledge improves.

So many individuals, both ACS members and outside reviewers, have contributed to the organization of this report that it would be impossible to mention them all by name. We are none the less grateful for the time and effort they took to ensure the completion of this report.

We would like to acknowledge the work of David Freeman (University of Maryland) and Wendell Phillips (Campbell Soup Co.) who prepared the initial drafts of the guidelines. Review and helpful commentary were received from Lloyd Currie (National Bureau of Standards), William Donaldson (EPA), William Horowitz (FDA), Richard A. Hummel (Dow), Robert A. Jamieson (Procter & Gamble), Gary L. Jewett (Dow), Eugene Meier (EPA), and Melvin Rueppel (Monsanto).

Daniel MacDougall	Warren Crummett
Chairman	Chairman
Committee on Environmental Improvement	Subcommittee on Environmental Analytical Chemistry

INTRODUCTION

Decisions regarding the use of chemicals are being based increasingly on results from analyses that measure chemical compositions near or below levels measured by conventional techniques. These trace analyses are subject to numerous difficulties including interferences, instrumental noise, and uncontrolled gain or loss factors; these introduce uncertainties in the final results.

The analytical chemist, whose task is to identify and quantify trace organics, is faced with special difficulties. The large number of organic compounds in the environment present the first problem; second is that the composition of the sample may be influenced by uncontrollable variables such as the effects of weather, tide, or time, which can complicate measurement and evaluation. Gaps in technical knowledge are also a problem, as is sampling bias, undetected flaws in the measurement systems, and errors in human performance. Systematic errors in the analyses are a major source of difficulty (*1*). For all these reasons, accuracy is very difficult to achieve in environmental trace analyses; yet it is extremely important.

When a costly and time-consuming study contains uncertainties such as those mentioned above, the results can be confusing and even meaningless. A well-designed and carefully executed measurement process is the best approach to remedy such problems. Such an analytical process may be conceptually visualized as a system consisting of the following components: plan model; sampling; measurement; calibration standardization; quality assurance; statistical procedures; documentation. If these components are well designed and functional, the probability of producing data of requisite quality increases. Conversely, if any of them are weak, nonfunctional, or missing, the quality of the data is jeopardized.

This document presents guidelines based on good analytical practices, to assist analysts in obtaining data of requisite quality and to aid in the evaluation of the quality of reported data.

PLANNING

Measurements must be based on a well-considered plan or model. Such a plan or model is implicitly or explicitly involved in every measurement process. Data are generated for use in answering questions or to provide information from which conclusions may be drawn. If the model, i.e., the interrelations of the data to the problem, is faulty, the conclusions drawn will be faulty, regardless of the quality of the measurements. The plan model should set forth the basic assumptions, i.e., the information–conclusions relations, and establish the basic measurement specifications. While the plan model may and should be revised as dictated by further information, no measurement program should be undertaken until a plan model is established.

An adequate plan model will be based upon a thorough understanding of the measurement process and a reasonable hypothesis as to the nature of the physical problem under investigation. It is best devised as a cooperative effort by the analysts with intimate knowledge of the measurement techniques, the scientist who will use the data, and the statistician who will evaluate the data. The plan should specify the general aspects of the components of the measurement system and refer to protocols (written details of the plans and procedures) of each of these.

QUALITY ASSURANCE

A quality assurance program is an essential part of a sound analytical protocol. Quality assurance is not a new concept. However, there is a clear need for its wider use by analytical chemists engaged in trace measurement. A single investigator as well as a laboratory organization should use quality assurance to detect and correct problems and take every reasonable step needed to keep the measurement process reliable.

A quality assurance program should include among others, the following elements (*2*):

1. Maintenance of skilled personnel, written and validated methods, and properly constructed, equipped, and maintained laboratory facilities
2. Provision of representative samples and controls
3. Use of high-quality glassware, solvents, and other testing materials
4. Calibration, adjustment, and maintenance of equipment
5. Use of control samples and standard samples, with proper records
6. Directly observing the performance of certain critical tests
7. Review and critique of results
8. Tests of internal and external proficiency testing
9. Use of replicate samples
10. Comparison of replicate results with other laboratories
11. Response to user complaints
12. The monitoring of results

13. Correction of departures from standards of quality

These basic elements of quality assurance define the framework that written protocols, including all analytical procedures, must follow to obtain reliable results. When a number of laboratories are involved, a uniform program should result in the convergence of interlaboratory comparison studies. The methodology in every protocol has limits beyond which measurements are unacceptable or where an excessive number of false positives or false negatives are produced so that the measurements are unreliable. When this happens, the measurement process becomes less than practicable (*3*). Therefore, it is important that the limits of reliability be specified as part of the protocol.

Laboratories must frequently rely on internal and external standards. The profession of analytical chemistry is being challenged by analytical problems of increasing difficulty. Critical assessment of the analytical system is needed to produce and validate analytical chemical measurements. Accurate data are far more likely to be obtained when supported by the competent use of external standards and when the laboratory has demonstrated close agreement with acceptable levels of accuracy, as determined by participation in interlaboratory comparisons.

SAMPLING

The quality of analytical data is critically dependent on the validity of the sample and the soundness of the sampling program.

The purpose of sampling is to obtain specimens that represent a larger population being studied. All aspects of a sampling program should be planned and documented in detail, and the expected relationship of the sampling protocol to the results should be made explicit. A sampling program should include reasons for choosing sampling sites, the timing of sample acquisition, and the accepted level of fluctuations due to heterogeneity. A detailed description of sampling sites and procedures is necessary and should include methodology, labeling, container preparation, field blank preparation, storage, and pretreatment procedures. An acceptable sampling program should *at least* include the following:

1. A proper statistical design which takes into account the goals of the studies and its certainties and uncertainties
2. Instructions for sample collection, labeling, preservation, and transport to the analytical facility
3. Training of personnel in the sampling techniques and procedures specified

The sample container and storage procedures must be consistent with the stability of the substance to be analyzed. It must be demonstrated that these do not alter the composition of the sample in a way that would affect the concentration of the chemical (analyte) being determined. Special transportation procedures (such as under refrigeration, exclusion of light) may need to be specified. Detailed guidelines for sampling design in some special situations are available (*4–7*).

SAMPLING REQUIREMENTS

In general, the number of samples and the quality of the sampling procedure must be controlled to characterize the sample and enhance reliability in the final results. If the sampling process is not otherwise specified, an investigator should decide what error level is tolerable, decide what confidence level is appropriate, and estimate the minimum number of samples required for confidence limits that satisfy the requirements of the measurement problem. Several approaches for defining the number of samples may be used.

Because environmental samples are ordinarily grossly heterogeneous, a large number must be analyzed to obtain meaningful data. The number of individual samples that need to be analyzed will depend on the kind of information required by the plan model. If an average compositional value is required, a large number of randomly selected samples may need to be obtained, combined, and blended to achieve a reasonably homogeneous composite of which subsamples may be analyzed. If composition profiles or the variability of the sample population is desired, many samples may need to be individually measured in replicate.

A statistical approach to sampling is possible when the standard deviation of the individual samples is known in advance or can be reasonably estimated. An example, described by Walpole and Myers (*8*), uses the expression

$$N = (Z\sigma_n/E)^2 \quad (1)$$

where N = number of samples, Z = constant (standard – normal) from tables (see ref 9, Table 2, p T-3), σ_n = standard deviation of individuals, and E = tolerable error in estimate of mean for characteristic measured.

For illustration, assume that the samples to be measured are expected to have a mean concentration of approximately 0.1 ppm with standard deviation of 0.05 ppm and that the tolerable error in the stated value of the mean at the 95% confidence level (Z = 1.96) is not to exceed 20% (0.02 ppm). A further assumption is made that the measurement error is small in comparison with the measured values and can be neglected in the calculation. With the above values, the number of samples required will be

$$N = [(1.96 \times 0.05)/0.02]^2 = 24 \quad (2)$$

One could either analyze 24 individual samples or combine them and analyze a homogenized composite. However, the composite would not give any information on the variability of the individual samples (valuable for checking the sampling strategy used) nor prove that a sufficiently homogeneous sample had been produced. The latter would require the analysis of a sufficient number of subsamples (seven is suggested).

Unfortunately, environmental trace analysis is often done where the standard deviation of the individual samples is not known in advance and where the measurement error cannot be predicted nor can it be assumed to be negligible. In this case, the measured values can be used to calculate an overall standard deviation, σ_o, which is related to the standard deviation of measurement, σ_m, and the standard deviation of individuals, σ_n, by the expression

$$\sigma_o^2 = \sigma_m^2 + \sigma_n^2 \quad (3)$$

An estimate of σ_m can be obtained by a pooling process, using the differences in the measured values of duplicate homogenized samples (see ref 10, p 316). Then the standard deviation of the individual samples, σ_n, can be calculated. Unless such calculations are based on a large number of measurements (at least seven), the standard deviation(s) may be significantly underestimated. In this case, the appropriate value of the Student's t test (ref 9, Table 4, pT-5) should be used and t values should be substituted for Z in eq 1 and similar expressions.

Equation 1 may also be used to estimate the number of replicate measurements, n, required on a homogeneous sample to achieve a mean value within a given confidence interval, E. In this case, σ represents the standard deviation of measurement. The following transposition of the equation

$$E = Z\sigma_m/n^{1/2} \quad (4)$$

may be used to calculate the confidence interval, $\bar{x} \pm E$, for the mean of n measurements.

When the data needed to calculate the minimum number of samples (N) are not available at the time of sampling, empirical approaches may need to be followed. In this case,

the *N–N–N* rule (*11*) is recommended as a helpful guideline. This means that equal numbers (*N*) of field samples, field blanks, and spiked blanks are to be analyzed along with the calibrating standards and controls. The rule was first used in U.S. Department of Agriculture pesticide residue studies as the 10–10–10 rule. This meant that a residue study required analysis of a minimum of 10 samples, 10 field blanks, and 10 spiked field blanks. A 7–7–7 rule is currently used by the U.S. Environmental Protection Agency in the analysis of water and waste water samples (*12*). Field blanks (sometimes called "control studies" in agricultural investigations) of environmental samples are often not readily available, but every effort should be made to obtain environmental blank samples that are believed to contain the analyte at levels below the limit of detection of the analytical method. In certain circumstances, a simulated or synthetic field blank is the only alternative.

SAMPLE PRETREATMENT

After a sample has been received, the analytical protocol may call for sample homogenization prior to subsampling. Special care should be taken regarding the relabeling, cataloging, inventory control, and storage processes. Special care must also be taken to prevent changes in the analyte concentration and the introduction of potentially interfering substances.

Sample pretreatment may involve physical operations such as sieving, blending, crushing, and/or drying and chemical treatments such as dissolution, addition of preservatives, standards, and other materials. These physical and chemical treatments must be documented in sufficient detail to provide a complete record of the sample history, so that another worker can exactly duplicate the treatment used.

MEASUREMENT PROCEDURE

Measurements should be made by using properly tested and documented procedures. Furthermore, every laboratory and analyst must do sufficient preliminary tests, using the methodology and typical samples, to demonstrate competence in the measurement procedure. Procedures that utilize controls and calibration steps to prevent (or expose) random and systematic error and procedures that provide high recovery, minimum contamination and acceptable precision are recommended. Error factors which need to be avoided include contamination by (or contaminants in) container, reagent, equipment, atmosphere, and added internal standards. The number of steps in the procedure should be kept to a minimum in order to reduce the possibility of error. The use of closed systems and as few manual operations as possible are recommended procedures in trace analytical methodology. Documentation must be sufficiently detailed to permit duplication by another qualified analyst.

CALIBRATION AND STANDARDIZATION

Calibration is the checking of physical measurements against accepted standards, including measurements of time, temperature, mass, volume, electrical units, and others. Standardization is defined (*13*) as determining the response function $S = g(C)$ where S is the measured net signal which is a function of "g" of the given analyte concentration (C). Where possible, one should carry out regression analysis consistent with a linear treatment of the function. For this purpose at least five different concentrations of the calibration standards should be measured in triplicate. The concentrations chosen should bracket the expected concentration of the analyte in the field sample. Standardization includes determining the signal response of an instrument when tested with varied amounts or concentrations of analyte in a suitable solvent or sample matrix. The instrumental response is the function $g(C)$ while the graphed version is known as the calibration curve. The standardization must be done under the same conditions as those that will exist during the measurement process, and this must be specified. This includes all of the assumptions concerning the calculation, such as assumed linearity, interpolations, and extrapolations. If testing has shown a reagent or condition that must be controlled, then the corresponding restrictions concerning the reagent or condition must be included as a part of the overall standardization process.

METHOD TESTING AND MONITORING

Precise measurements depend largely on the proper use of good laboratory practices, proven methodology, and low noise instrumentation. Accuracy is supported by the use of Standard Reference Materials, and participation in interlaboratory comparison activities. The source, composition, and purity of standards must be established; such information is available from the U.S. National Bureau of Standards, U.S. Environmental Protection Agency, U.S. Food and Drug Administration, U.S. Department of Agriculture, and various commercial sources. The purity of a standard should be measured and documented, possible interferences should be determined, and the possible time dependence of the composition of a standard should be measured to determine whether or not any change has occurred during the analytical study.

Some degree of inaccuracy occurs in every measurement, so a strategy is needed which minimizes errors. Elements of this strategy include minimizing the complexity of the analytical procedure and evaluating every experimental variable as a possible source of error, thus reducing the opportunities for error. Errors may arise when the measurement process is unexpectedly sensitive to seemingly small changes in operation (*14*). Examples of this are the possible sensitivity of a method to laboratory temperature, humidity, or the concentration of certain additives to the liquid carrier in liquid chromatography. The purity of reagents—solvents, chemicals, and gases used in the analysis—must be known or measured as a part of the quality assurance program applied to reagent acquisition and purification. The reagent blank is a useful indicator but it should not substitute for this critical information. Youden devised a plan for examining the effect of changing more than one variable at a time, known as ruggedness testing (*15*, *16*). Each selected analytical method should be exhaustively tested for error sources before it is used, so that corresponding controls can be specified if necessary. Careful adherence to this approach should permit a reasonable assessment of the precision and reliability of the data to be attained.

Any proven method should be retested during the measurement process by periodic analyses of the blanks, standards, and "spiked" samples. A knowledge of possible interferences needs to be established beforehand (*16*). If method testing activities reveal a critical condition or reagent, steps should be taken to monitor the corrective conditions intended to prevent anomalous results.

DEFINITION OF THE DATA SET

The proper conduct of trace analytical measurements requires that data be obtained from the following six sample categories, whenever possible: (a) calibration standards, (b) field blanks, (c) spiked field blanks, (d) spiked laboratory blanks, (e) working standards, (f) field samples. It may be helpful to add a set of spiked field samples, as well.

The use of these samples helps validate the measurement process. If the calibration standards are stable and recovery is reproducible, then more time can be devoted to measuring the field samples and field blanks. (The use of controls tends to be omitted from radioactivity determinations; when the

analyte signal approaches the background, equal time is usually allotted to sample and background measurements.) Since multiple manipulations tend to increase measurement variability, performance testing is needed to monitor the recovery and variability in measuring knowns and blanks. The performance tests are based on the use of working standards whose values are well-known and, where possible, on Standard Reference Materials. The frequency and order for measuring a sequence of blanks, controls, and field samples should depend on each protocol or measurement situation (*17*). In carrying out recovery studies the analyst should know that the analyte added to a blank sample may behave differently (typically, showing higher recovery) from the analyte in the field sample.

PERFORMANCE TESTING AND DATA HANDLING

Electronic data handling, data reduction, and data storage systems are important parts of many modern analytical systems. They greatly facilitate data handling and help control errors due to misreading, faulty transcription, or miscalculations. However, the performance of the data system must be tested regularly to ensure that it is working correctly. This should be done periodically by using known data that have already been calculated. These tests must have sufficient accuracy and precision to provide a reliable test of the data handling system.

MEASUREMENT VARIABILITY

Every analytical system contains sources of inaccuracy and imprecision, both of which have variable components. Systematic errors affect the accuracy by giving a net bias plus fluctuations that may introduce an apparent random contribution. Random errors are also present. Such variations arise due to weighing uncertainties, aliquoting variability, residual sample heterogeneity, instrumental noise, and other sources.

There are two basic indicators of measurement quality: precision and accuracy. When experimental measurements are first carried out by use of reliable working standards, excessive measurement variability provides a good reason to search for uncontrolled systematic errors. When evaluating results by validated methods, the final estimate of precision will usually rely on the assumption that all practicable steps have been taken to control (i.e., suppress, eliminate, or compensate) the systematic errors. The remaining fluctuations are considered random and will determine the experimental precision.

CONCENTRATION VARIABILITY

The following discussion is consistent with well-established criteria for assessing the measured signal response, its variability, and the statistical significance of the numerical results.

When sample measurements are made, the observed signal S_t is the sum of the instrumental response S_x due to the presence of the analyte x in the sample, plus a response signal S_b due to the background or blank contribution. Thus

$$S_t = S_x + S_b \tag{5}$$

When a separate measurement of the blanks is made, one obtains an estimate of S_b. The net analyte signal can be estimated from the difference

$$S_x = S_t - S_b \tag{6}$$

The net analyte concentration C_x is found by using the response function, $S_x = g(C_x)$, or by graphic interpolation of the calibration curve as previously defined (common practice calls for plotting some function of S_x which is linear with C_x).

The relative variability of analytical measurements characteristically increases as the analyte concentration diminishes. There are three regions of reliability which, in descending order of reliability, are herein defined as the levels of determination, detection, and uncertain detection. The absolute signal variability (σ) is defined by the standard deviation in the estimated net signal (S_x). This quantity should be based on sound knowledge of the particular measurement process being used with and obtained from at least 10 observations, if feasible. For any signal variability, σ_s, the corresponding concentration variability, σ_c, can be calculated from the response function ($S = g(C)$). In the particular case of a linear dependence of S_x upon C_x

$$\frac{\sigma_{C_x}}{C_x} = \frac{\sigma_{S_x}}{S_x} \tag{7}$$

These ratios determine the relative standard deviation. Since the net signal is obtained by subtraction, the propagation of errors leads to the following expression for the linear case:

$$\frac{\sigma}{S_x} = \frac{\sigma_{C_x}}{C_x} = \frac{(\sigma_t^2 + \sigma_b^2)^{1/2}}{S_x} \tag{8}$$

Where the key term σ refers to the error σ_{S_x} in the net analyte signal. Furthermore, the uncertainty in the mean of n measurements is inversely proportional to the square root of the total number of measurements. This applies to uncertainty of the value of C_x and to the standard deviation of the blank σ_b which plays a critical role in deciding the level of validity that has been reached in trace analysis.

LIMIT OF DETECTION (LOD)

The limit of detection is the lowest concentration of an analyte that the analytical process can reliably detect. Method sensitivity is defined (*18*) as the ratio of change in the instrument response to the change in analyte concentration, i.e., the slope of the calibration curve should not be confused with limit of detection. The LOD in most instrumental methods is based on the relationship between the gross analyte signal S_t, the field blank S_b, and the variability in the field blank σ_b. The limit of detection has been variously defined (*19–21*) by the extent to which the gross signal exceeds S_b.

$$S_t - S_b \geq K_d\sigma \tag{9}$$

If field blanks are not available or if a single sample is being analyzed for which there is no field blank data, then the limit of detection is based on the peak to peak noise ($\sigma = \sigma_n$) measured on the base line close to the actual or expected analyte peak. It is recommended in these guidelines that detection should be based on a minimal value for K_d of 3. Thus, the limit of detection is located at 3σ above the gross blank signal, S_b. This is illustrated in Figure 1 which depicts the region for detection of analyte in the gross signal as $S_t \geq S_b + 3\sigma$ and similarly for the net signal as $S_x \geq S_d$.

The corresponding limit of detection expressed in concentration units is derived from the calibration function. To illustrate this, consider a simple linear calibration function where $S_x = aC_x$. The lower limit to detection, expressed in concentration units, would then be obtained from the net signal C_x (LOD) $\geq S_d/a$. While a value of $K_d = 3$ is considered minimal, higher values may be required since $K_d = 3$ implies definite risks of 7% for false positive (concluding the analyte is present when it is absent) and false negative (the reverse) decisions. The question of risk levels associated with various K_d values is discussed more completely by Kaiser (*20*) and Currie (*21*). To illustrate, Kaiser has used a more conservative criterion for the purity of a compound with $K_d = 6$, thereby

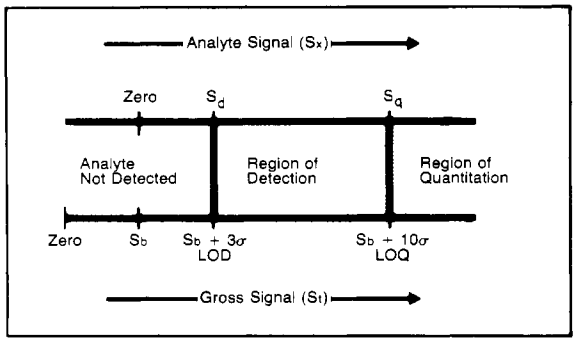

Figure 1. The limit of detection (LOD) is located 3σ above the measured average blank. The limit of quantitation is 10σ above the blank. These are the minimal criteria recommended by these guidelines.

decreasing risks of false positives or negatives.

If individual field samples are measured for screening purposes, further replication may be required by the protocol. For illustration, initial positive results on actual field samples can be validated by repeated analyses of subsamples from the same field samples. Replication leads directly to improved precision as illustrated by eq 4. The precision of the mean value is thus related by inverse proportion to the square root of the total number of measurements. Agreement between replicate analyses above the LOD increases the likelihood that the analyte has truly been detected. However, the result is not validated until one or more independent methods provide consistent results.

Accuracy in the determination of the value of σ improves as the number of measurements increases. If the replication is limited, it is important to increase the apparent or estimated value of σ to prevent serious underestimation of its value. For this purpose, the use of the distribution based on Student's t test is required. It should be noted that the measured standard deviation converges on the "true value" (σ) only after a large number of measurements.

LIMIT OF QUANTITATION (LOQ)

The numerical significance of the apparent analyte concentration increases as the analyte signal increases above the LOD. As a minimum criterion, the region for quantitation should be clearly above the limit of detection. Details of the chosen criterion should be consistent with the purpose of the plan model whether it be a survey measurement, screen, quality control, legally required monitoring, or evidence for possible violation of a legal limit. The following definitions are useful. The limit of quantitation (LOQ) is located above the measured average blank S_b by the following definition:

$$S_t - S_b \geq K_q\sigma \qquad (10)$$

It is recommended that the minimum value be $K_q = 10$. This is illustrated in Figure 1 where the region for analyte quantitation is expressed in terms of the gross signal, $S_t \geq (S_b + 10\sigma)$, and in terms of the analyte signal, $S_x \geq 10\sigma$. The corresponding expression in terms of concentration would be obtained by using the calibration function. Following the previous example used for the LOD, $C_x(\text{LOQ}) = S_q/a$ (for $S_x = aC_x$).

The preceding definitions of LOD and LOQ have now been stated mathematically and illustrated in Figure 1. The combined definitions can now be seen to define the region of detection $S_q > S_x \geq S_d$ and the region of quantitations $S_x \geq S_q$. These regions are shown in Figure 1 and in Table I.

Signals below 3σ should be reported as not detected (ND) with the limit of detection given in parentheses. Signals in the region of detection should be measured and reported as numbers with the limit of detection given in parentheses. The symbols "T" or "tr" for trace amounts have occasionally been used to refer to results in a region of high uncertainty between LOD and LOQ. It is recommended that this latter practice be abandoned.

Table I. Regions of Analyte Measurement

analyte signal (S_x)	recommended inference
$<3\sigma$	analyte not detected
3σ to 10σ	region of detection
$>10\sigma$	region of quantitation

Environmental measurements are often accompanied by appreciably large relative errors, σ_t/S_t, of 0.1–0.25 or even greater. Further, it is often found that gas chromatographic or coupled mass spectrometric determinations have σ_t varying proportionally with S_t. In these and other cases the level of error must be reduced. The recommended approach is to increase the level of replication because the relative standard deviation unlike the mean deviation may not be small enough to meet the condition for quantitation.

In many cases, a field blank sample may be difficult to define or obtain experimentally. If the field blank (actual or simulated) is not properly defined in the protocol, then the results are invalid. If individual field samples are measured for screening purposes, further analysis may be required by the protocol, again by stating the requirements for replication before the measurements are actually carried out. For illustration, initial measurement screens on single or duplicate field samples can be validated by repeat analysis of subsamples from the same field samples. Agreement between the analyses will obviously increase the reliability of the measurement. However, the quantitative result is unconfirmed until one or more independent methods provide the confirmation, as will be discussed shortly.

QUALITATIVE CONFIRMATION OF VALIDATED MEASUREMENTS

It is necessary to verify that the qualitative identity of the measured apparent compound in a sample is the same as the analyte which is known in the reference chemical standard. This must be based on an analytical principle and/or analytical conditions different from those used in the initial analytical method. This confirmation procedure should be highly selective and it should refer to an unambiguous property that is characteristic of the analyte. To illustrate, one GC/MS method may be validated by another GC/MS method which differs in the chromatographic conditions, methods of ionization, and/or methods of detection.

ADDITIONAL RISKS IN LOW RECOVERY METHODS

The recovery of a method is derived from the measurement of "spiked blanks". These may be controls or simulated field samples containing varied known added concentrations, C, of the analyte. The recovery is determined by using

$$\%\ \text{recovery} = \frac{C(\text{found})}{C(\text{added})} \times 100 \qquad (11)$$

where C(found) is based on the net analyte signal for the "spiked" blank. High recoveries leave intrinsically little room for variability in the recovery itself. Conversely, as the recovery falls, the measurement process becomes more dependent on the knowledge of the precision of the actual recovery at that concentration. In carrying out recovery studies, the analyst should know that analyte added to a blank sample may behave differently (typically, showing higher recovery) from analyte in the same sample. In this case, the method of standard addition tends to lead toward erroneously low values. Whenever possible, testing should include experiments on homogeneous working standards containing known

amounts of naturally incorporated analyte. Unfortunately, the frequent lack of such specified samples or standards is an important limitation in modern trace organic analysis.

Alternatively, the use of low recovery methods may be satisfactorily applied in the "region of quantitation" only if the accuracy and precision are established at a specified confidence limit. Under these conditions the measurements of field samples, when corrected for recovery, can accurately indicate the true analyte concentration. Recoveries of less than 60% are usually considered unreliable due to the difficulty of obtaining its value accurately.

VERIFICATION

The reliability and acceptability of environmental analytical information depend upon the rigorous completion of all the requirements stipulated in a well-defined protocol. Such protocols will describe the documentation requirements of the study including sampling procedures, measurements, confirmation, and validation. In addition, all results should be critically reviewed. If questions arise during the review, additional confirmatory tests should be conducted, including the use of methods other than those previously applied. In situations where large numbers of samples are analyzed, using widely accepted and well-documented analytical systems, unusual high results, or unexpectedly low ones, should be validated by a repeat analysis of a duplicate subsample using the same method and a third subsample analyzed by a different analytical method. Agreement of the three results indicate that the analyte is present.

Confidence in the measurement process can be strengthened considerably by interlaboratory comparisons. This is one of the most effective elements of a quality assurance plan. Successful collaborative testing (22) should be required for all analytical methods specified for use in major government decision-making processes. In some laboratory studies, especially those involving the use of a protocol for the first time, the interlaboratory comparison of careful homogenized and preserved subsamples may indicate serious discrepancies in the initial round due to undetected errors. Efficient planning can reduce the effort needed to improve accuracy. Currie and DeVoe have recommended (*23*) that Youden's correlation technique (*24*) should be used in interlaboratory comparisons in order to distinguish between random errors, laboratory bias, and erratic blunders.

EXPRESSING RESULTS

Measurement results should be expressed so that their meaning is not distorted by the reporting process. The significance of any number is indicated by the way the number is stated. The data are rounded off so that only the last figure reported is in doubt. However, this method of reporting does not clearly indicate the confidence intended for the data; hence dependence on significant figures for indicating confidence level is discouraged.

Data obtained according to these "guidelines" will permit statistical treatment, and this practice is recommended. Expressed results should include the average measured value, and the uncertainty should be made explicit by including the number of measurements and the standard deviation. In particular, the data should clearly show the relationship between individual sample values, blanks, recoveries, averages, variabilities, and the relative standard deviation. The average value should be written with the standard deviation of the mean, and details should be presented showing how the averaging process accounts for sample heterogeneity as well as the observed imprecision among the replicate measurements of homogenized subsamples (*25*). (The experimental design should anticipate possible problems in dealing with heterogeneity as well as reporting its effect.) It is an explicit requirement that final results include scientifically based estimates of the bounds for systematic error.

DOCUMENTATION

The reporting process should provide sufficient information to support claims made for the results and an estimate of uncertainty. Any new methodology used should be described in detail, including the results of exhaustive tests of the methodology. The use of existing methodology can be cited by reference to published literature, provided that the published description is complete and adequate and that any modifications have been fully tested and reported.

It is equally important to illustrate with an example of how the concentration of the analyte in analyzed samples was calculated. Raw data for each sample along with the results of reagent blanks, control, and "spiked" samples should be suitably identified. Raw data should include pertinent information such as complete sample specification, transfers and movement, sample number, initial sample weight, extraction volume, final weight and volume analyzed, instrument response, chromatograms or instrument response, weight of compound found, and concentration of sample. All laboratory records or suitable copies should be permanently filed for future reference.

INTERFERENCES REPORTED

A complete report should consider the possible interferences which can arise at any stage in the analytical process. Reagent blanks, cleanup steps, and effects of additives, light, and time should be evaluated as possible interferences. Compounds closely related to the analyte should be evaluated for possible negative or positive interference effects. When interferences are studied, the results should be tabulated; when they are not, an explanation should be given. The use of detailed testing to reveal sensitivity to interferences before the adoption of a method is absolutely essential to ensure reliability.

CONCLUSION

Accurate environmental chemical analysis cannot be based solely on the ability of an instrument to respond to small concentrations in a sample. Trace analysis requires properly skilled and properly trained analytical chemists. Alertness and discipline must be applied. Thus, modern trace analysis is built upon three interwoven strategies: the development of sensitive, specific and validated analytical methods; the use of protocols that describe the details of the measurement process and the sampling procedures; and the use of quality assurance procedures to monitor the quality of the data as it is developed. No laboratory or individual should undertake to provide environmental analytical data unless they are aware of the importance of using a comprehensive and systematic approach and are committed to making the necessary effort to prepare reliable measurements. Of utmost importance is the validation process which determines whether the measurement process is sound. Today it is necessary to communicate the results of such measurements in a way that permits an open inspection of their intrinsic weaknesses and strengths.

LITERATURE CITED

(1) Hilpert, L. R.; May, W. E.; Wise, S. A.; Chesler, S. N.; Hertz, H. S. *Anal. Chem.* **1978**, *50*, 458.
(2) Inhorn, S. L., Ed. "Quality Assurance Practices for Health Laboratories"; American Public Health Service: New York, 1978, p 199
(3) Horwitz, William *Anal. Chem.* **1979**, *51*, 741A.
(4) "EPA Proposed Guidelines for Registering Pesticides in the U.S.", FR 40, No. 123; Environmental Protection Agency: Washington, DC, June 25, 1975; pp 26802–26928.
(5) "EPA Proposed Guidelines for Registering Pesticides in the U.S.", FR 43, No. 132; Environmental Protection Agency: Washington, DC, July 10, 1978; pp 29696–29741.
(6) deVera, E. R.; Simmons, B. P.; Stevens, R. D.; Storm, D. L. "Samplers and Sampling Procedures for Hazardous Waste Streams"; EPA-600/2-80-018; Environmental Protection Agency: Washington, DC, Jan 1980.
(7) Scheide, E. P.; Filliben, J. J.; Taylor, J. K. "Survey of the Occurrence of Mercury, Lead, and Cadmium in the Washington D.C. Area"; NBSIR

78-1428; National Bureau of Standards: Washington, DC, Sept 1977.
(8) Walpole, R. E.; Myers, R. "Probability and Statistics for Engineers and Scientists"; Macmillan: New York, 1972; p 190.
(9) Natrella, M. G. "Experimental Statistics"; *NBS Handb. (U.S.)*, **1963**, *No. 91*.
(10) Ku, H., Ed. "Precision Measurement and Calibration"; *NBS Spec. Publ., No. 300*, Vol. 1.
(11) Harris, T. H.; Cummings, J. G. *Residue Rev.* **1964**, *6*, 104.
(12) "Handbook for Analytical Quality Control in Water and Wastewater Laboratories"; U.S. Environmental Protection Agency, Environmental Monitoring and Support Laboratory: Cincinnati, OH, March 1979; p 10-6.
(13) DeVoe, J. R., Ed. *ACS Symp. Ser.* **1977**, *No. 63*.
(14) Wernimont, G. *ACS Symp. Ser.* **1977**, *No. 63*, Chapter 1.
(15) Youden, W. J. *Assoc. Off. Anal. Chem.* **1963**, *46*, 56.
(16) Youden, W. J.; Steiner, E. H. "Statistical Manual of the Association of Official Analytical Chemists"; Association of Official Analytical Chemists: Washington, DC, 1975.
(17) "Sixth Materials Research Symposium in 1973"; *NBS Spec. Publ. (U.S.)* **1975**, *No. 408*, 806.
(18) "FDA Proposed Criteria and Procedures for Evaluating Assays for Carcinogenic Residues", FR 44, No. 55; Food and Drug Administration: Washington, DC, March 20, 1979; pp 17070-17114.
(19) "Appendices on Tentative Nomenclature, Symbols, Units and Standards, II. Terms and Symbols Related to Analytical Functions and their Figures of Merit—Number 28"; International Union of Pure and Applied Chemistry, Oxford, U.K., 1972.
(20) Kaiser, H. H. *Anal. Chem.* **1970**, *42*, 26A.
(21) Currie, L. A. *Anal. Chem.* **1968**, *40*, 586.
(22) Youden, W. J.; Steiner, E. H. "Statistical Manual of the Association of Official Analytical Chemists"; Association of Official Analytical Chemists: Washington, DC, 1975, Chapter 4.
(23) Currie, L. A.; DeVoe, J. R. *ACS Symp. Ser.* **1977**, *No. 63*, Chapter 3.
(24) Youden, W. J. *Anal. Chem.* **1960**, 32, 23A.
(25) Hertz, H. S.; Chesler, S. N. Ed. *NBS Spec. Publ. (U.S.)* **1979**, No. 519. See articles on sampling by H. Ku, p 1; B. R. Appel et al., p 121; H. G. Eaton et al., p 213; W. Horowitz and J. W. Howard, p 231; and H. G. Lento, p 243.

David Freeman acknowledges support from the National Science Foundation for his work in connection with this project.

Distribution Phenomena of Mobile-Phase Components and Determination of Dead Volume in Reversed-Phase Liquid Chromatography

R. M. McCormick and B. L. Karger*

Institute of Chemical Analysis and Department of Chemistry, Northeastern University, Boston, Massachusetts 02115

The elution behaviors of three common reversed-phase organic modifiers were examined from 0 to 100% modifier in aqueous mobile phases. Elution of modifier and deuterated modifier enriched samples is complicated by retention and isotherm curvature. The former species yields a concentration pulse in the chromatogram while the latter produces two bands—a concentration pulse and a band containing the deuterated species. Injection of water-enriched samples results in vacancy peaks with the same elution behavior as corresponding modifier-enriched samples. Injection of D_2O-enriched samples yields two bands—a vacancy peak and the D_2O-rich band. The elution volume of D_2O provides a good estimate of the column dead volume. Modifier distribution isotherms measured by several procedures show that significant quantities of organic solvent are extracted into the stationary phase, the amount being dependent on the solvent strength of the modifier and its concentration in the mobile phase.

At present, the majority of high-performance liquid chromatographic (HPLC) separations are accomplished with use of chemically bonded reversed-phase packings. The predominant mechanism of retention in reversed-phase liquid chromatography (RPLC) is hydrophobic expulsion of a solute from a mixed aqueous organic mobile phase (*1*, 2). While hydrophobic phenomena are generally associated with nonpolar solute selectivity, it is now also recognized that selectivity based on polar group differences of solute molecules can also be very significant in RPLC (*3-5*). The organic solvent or modifier (e.g., methanol (MeOH), acetonitrile (ACN), or tetrahydrofuran (THF)) has been shown to play a particularly important role in this behavior (*3*, *6*), and thus control of mobile-phase composition (e.g., ternary phases (*3*, *4*)) provides a powerful means of manipulating separation.

To work successfully with mobile-phase phenomena in the achievement of separation, one needs to understand in more detail the retention process in RPLC. At the heart of this understanding, a better description of the stationary phase is required, as well as a more complete consideration of the distribution phenomena of all components in the reversed-phase column.

RPLC retention is more complex than simple solute interaction with the bonded *n*-alkyl chains. Besides the influence of unreacted and accessible silanol groups, the organic modifier, as first proposed by Knox and Pryde (*7*), is enriched in the stationary phase, and this extracted solvent will also interact with solute species. Scott and Kucera (*8*, *9*), Westerlund and Theodorsen (*10*), and Tilly-Melin et al. (*11*) provided experimental proof of the selective concentration of the modifier into the bonded phase.

The organic solvent is expected to distribute between the mobile and stationary phase in RPLC since the modifier, as well as the solute, is subject to hydrophobic expulsion from the aqueous mobile phase. The quantity of extracted modifier will be dependent upon its solvent strength, and thus THF, for example, should extract to a much larger extent than MeOH. The composition and character of the stationary phase will vary, depending upon the composition of the mobile phase. Hence, in order to understand retention in RPLC, the distribution process of the organic modifier must be considered.

The purpose of this paper is to examine in detail the distribution of organic modifiers in reversed-phase columns. Consequences of this distribution process with respect to retention of various species identical or related to the mobile-phase components (e.g., deuterated solvents) will also be examined. Since this distribution is an equilibrium process and since the organic modifier is a constituent of the mobile phase, the laws of finite concentration chromatography are applicable to the organic solvent. On the basis of the information developed in this work, a second paper (*12*) will deal with retention phenomena of solute species (i.e., nonconstituents of the mobile phase in RPLC).

0003-2700/80/0352-2249$01.00/0