A NIOSH Technical Report

Guidelines for Air Sampling and Analytical Method Development and Evaluation

U.S. DEPARTMENT OF HEALTH AND HUMAN SERVICES
Public Health Service
Centers for Disease Control and Prevention
National Institute for Occupational Safety and Health

CDC
CENTERS FOR DISEASE CONTROL AND PREVENTION

Guidelines for Air Sampling and Analytical Method Development and Evaluation

Eugene R. Kennedy, Ph.D., Thomas J. Fischbach, Ruiguang Song, Ph.D.
Peter M. Eller, Ph.D. and Stanley A. Shulman, Ph.D.

U.S. DEPARTMENT OF HEALTH AND HUMAN SERVICES
Public Health Service
Centers for Disease Control and Prevention
National Institute for Occupational Safety and Health
Division of Physical Sciences and Engineering
Cincinnati, Ohio 45226

May 1995

DISCLAIMER

Mention of company names or products does not constitute endorsement by the National Institute for Occupational Safety and Health.

DHHS (NIOSH) Publication NO. 95-117

ABSTRACT

The Occupational Safety and Health Act of 1970 (Public Law 91-596) charged the National Institute for Occupational Safety and Health (NIOSH) with the responsibility for the development and evaluation of sampling and analytical methods for workplace compliance determinations. Under that charge, NIOSH and the Occupational Safety and Health Administration jointly undertook the evaluation of sampling and analytical methods for airborne contaminants by contract in 1974 to determine if methods met the criterion to produce a result that falls within 25% of the true value 95 times out of 100. This guideline document further expands the experimental protocol used during this initial methods development and evaluation research.

The experiments listed in this document include determination of analytical recovery from the sampler, sampler capacity, storage stability of samples, and effect of environmental factors. Also included are evaluation criteria for the experiments and details for the calculation of bias, precision and accuracy.

ACKNOWLEDGEMENTS

The authors of this document, Eugene R. Kennedy, Thomas J. Fischbach, Ruiguang Song, Peter M. Eller and Stanley Shulman acknowledge that this document is the result of a process that began in 1974 with the joint NIOSH/OSHA Standards Completion Program. The foundations which were established by this program defined the AC and the evaluation procedures developed to verify method performance. These are still key elements in this document. As the state of the art has evolved, we have attempted to include this information in our evaluation procedures. Where these additions have been made, we have attempted to reference the source.

CONTENTS

I. INTRODUCTION

A. Background

The Occupational Safety and Health Act of 1970 (Public Law 91-596)[1] charged the National Institute for Occupational Safety and Health (NIOSH) with the responsibility for the development and evaluation of sampling and analytical methods for workplace compliance determinations. Under that charge, NIOSH and the Occupational Safety and Health Administration (OSHA) jointly undertook the evaluation of sampling and analytical methods for airborne contaminants by contract in 1974. During the course of this work, an experimental protocol was developed to define the evaluation criteria to be used for method evaluation.[2,3,4] For each method under consideration, the objective of this protocol was to determine if the method would provide results that were within ±25% of the (true) concentration 95% of the time.

During the course of methods development and evaluation research, the effects of certain environmental and experimental conditions on sampling and analytical method performance have been documented. The purpose of this guideline document is to further refine the original protocol for application to sampling and analytical method development and evaluation research with additional experiments added to more fully evaluate method performance. An experimental design for the evaluation of sampling and analytical methods has been suggested. If these experiments are not applicable for the method under study, then a revised experimental design should be prepared which is appropriate to fully evaluate the method. The assistance of a statistician may be required for the preparation of this design.

The objectives of this guideline document are: 1. To provide guidance and procedures to estimate the precision, bias, and accuracy of a sampling and analysis method: in the case of accuracy, the estimates include the single value which is the best descriptor of the accuracy, and a 90% confidence interval estimate[a]; and 2. To provide guidance and procedures to evaluate a method relative to the 25% accuracy criterion (or one specified by the user) in terms of one of three mutually exclusive possible conclusions: a) a definite positive conclusion that there is 95% confidence that the method achieves the accuracy criterion; b) a definite negative conclusion that there is 95% confidence that the method fails the accuracy criterion, i.e., that, at best, the method accuracy is worse than 25%; or c) that the evidence concerning whether the method does or does not fulfill the accuracy criterion is inconclusive and that further research is required to resolve the question.

The work described in this document can be summarized in five steps as follows: 1) selection of compounds for method development and evaluation; 2) development of the sampling and analytical method; 3) evaluation of the method; 4) preparation of a written version of the method; 5) preparation of a technical report on the development and evaluation.

B. Definitions

This section defines some terms that are used in the rest of this document. Many of these terms are quantities (e.g., bias) and the procedures for calculation are listed in III. Method Evaluation.

1. *Accuracy* The ability of a method to determine the "true" concentration of the environment sampled. Accuracy describes the closeness of a typical measurement to the quantity measured although it is defined and expressed in terms of the relative discrepancy of a typical

---

[a] Unless explicitly stated otherwise, all confidence interval estimates used in these Guidelines are two-sided intervals.

measurement from the quantity measured. The term "inaccuracy" has also been used interchangeably with the term "accuracy" in the literature.[5] In this document, only the term "accuracy" will be used. Accuracy can be a characteristic of a method when measurements follow a statistical distribution, such as the normal[a] distribution. The special sense of accuracy for a method is embodied in the following definition and criterion:

> The accuracy of a method is the theoretical maximum error of measurement, expressed as the proportion or percentage of the amount being measured without regard for the direction of the error, that is achieved with 0.95 probability by the method.

The accuracy criterion (AC) used in the previous protocol[2,3,4] and in this document requires that a method give a result that is within $\pm$25% of the true concentration with a probability of 0.95 for an individual observation (i.e., that the accuracy of an acceptable method is no greater than 25%).

The region below a given curve in Figure 1 (page 25) shows all the combinations of relative standard deviations (S_{rT}) and biases (B) which would result in method accuracy at or better than the stated accuracy level for a given curve. Table I provides a representative set of the numerical values corresponding to the graph in Figure 1.

For a method to be accepted as fulfilling the AC, the data from the evaluation study must provide 95% confidence that the accuracy of the analytical method is at least at the level of the AC (25%). To obtain 95% confidence that the accuracy of a method satisfies the AC, the 95% confidence limit estimate of the accuracy (see Appendix 1) must be less than 25%. For a method to be rejected as definitely not fulfilling the AC, the 5% confidence limit estimate of the accuracy (see Appendix 1) must be greater than 25%. If neither of these conditions can be justified, the results are inconclusive and more research will be required to reach a definite acceptance or rejection of the method. Alternatively, the method can be rejected at this point if resources are not available for further research.

2. *Precision* The relative variability of measurements on replicate samples about the mean of the population of measurements, designated by σ, divided by the mean at a given concentration, designated by μ. The term "imprecision" has also been used interchangeably with the term "precision" in the literature.[5] In this document, only the term "precision" will be used. Precision is expressed by the relative standard deviation, denoted by S_{rT} (See Appendix 2), of a series of measurements. It reflects the ability of a method to replicate measurement results. The statistical definition of the value is given by:

$$(1)\quad S_{rT} = \sigma/\mu$$

These guidelines assume that the S_{rT} of the evaluated method is constant or homogeneous over all concentrations tested for the method evaluation[b]. This assumption should be tested using the procedures described in Appendix 2 (Bartlett's Test).

---

[a] The normal distribution assumption is used for several reasons. It is reasonable as the model for analytical errors -- which are measurement errors -- even though the environmental concentrations measured may be lognormal. Unpublished results for the methods studied in references 2 to 4 indicate that there is little empirical inconsistency with that assumption. Normal theory results are often applicable for other cases or as good first approximations. Moreover, aside from the S_r estimates, the analysis is "means-based." Finally, the authors' unpublished results show that relationships among the method accuracy, precision, and bias that follow from normal theory assumptions hold extremely well for several other distributions, e.g., lognormal, gamma, etc.

[b] This assumption does not imply that the relative standard deviations of methods are constant over all concentrations, only those selected for the study.

3. *Bias* The uncorrectable relative discrepancy between the mean of the distribution of measurements from a method and the true concentration being measured, T, expressed as a fraction. It is given by:

$$(2) \quad B = [(\mu/T) - 1]$$

Bias does not include correctable bias, such as recovery efficiency corrections. Acceptable methods must have an absolute bias no greater than 10%. A statistical test is described in Appendix 1.

These guidelines assume that the evaluated method bias is constant over all concentration levels tested[a]. This assumption should be tested using the procedures described in Appendix 1 for evaluating homogeneity of the bias.

4. *Limit Of Detection and Limit Of Quantitation* The Limit Of Detection (LOD) and Limit of Quantitation (LOQ) are defined in the current version of SOP 018[6] which appears in Appendix 3.

The LOD and LOQ should be considered as guidelines of method performance and should not be considered as absolute values. The LOD determined using SOP 018 indicates that the analyst knows with ≈99% confidence that instrumental signal at or greater than the LOD is due to the analyte.

The LOD, as well as the LOQ, may vary from laboratory to laboratory, analyst to analyst, instrument to instrument and day to day. Therefore, any determination of this value should be performed under the same conditions used for sample analysis and only reported with those analyses. When this value is reported in a sampling and analysis method, it should be stressed that this value is only an estimate of the expected performance of the method.

5. *Measurement Range* Concentration range of the analytical standards used for method evaluation.

6. *Evaluation Range* Range of generated concentrations over which the method was evaluated. For most analytes, this range covers concentrations from 0.1 to 2.0 times the Exposure Limit.[7] In some cases, this range may be extended to include 10 times the Exposure Limit.[8] In cases where an atmosphere of an analyte was not generated, the evaluation range can be calculated as the range of concentrations which would be equivalent to the amounts of analyte fortified onto the samplers for the evaluation experiments, based on typical sampling times and rates.

7. *Estimated Recovery* The ability to recover and determine an analyte placed in or on a sampler. Recovery efficiency is estimated by dividing the amount recovered from a sampler by the amount of analyte fortified in or onto the sampler. Acceptable analytical recovery has previously been defined as 75% or greater for six or more replicates at the Exposure Limit.[2,3]

8. *Interferences* Other compounds or conditions that are present with the analyte in the environment sampled can be an interference in the determination of the analyte. Potential interferences are listed below:

   a. Compounds that interfere with the resolution of the analyte during analysis.

   b. Compounds that interfere with efficient collection of the analyte during sampling.

---

[a] The assumption of a constant bias applies only to the range of concentrations tested for the evaluation study and not in general.

c. Compounds that interfere with the recovery of the analyte from the sample media.

d. Conditions that interfere with collection or analysis of the sample.

9. *Capacity (Breakthrough)* The maximum amount of analyte that can be collected on the sampling media without diminishing the ability of the sampling media to collect analyte as defined by sampling rate, concentration, and sampling time. This parameter also can be expressed in terms of time or total volume with the sampling rate range and concentration specified in each case.

   Criteria that have been used for defining this parameter include:

   a. Volume sampled (flow rate x time) until breakthrough of 5% of influent concentration through the sampler.[4]

   b. Volume sampled (flow rate x time) until mass found on the backup section of the sampler totaled 5% of the mass found on the front section of the sampler.[4]

10. *Sampling Rate or Uptake Rate* Volumetric rate that the air containing the analyte is taken into sampler. For vapor samplers this rate is further defined such that no breakthrough occurs at a set concentration for a defined period of time. For particulate samplers, the pressure drop of the sampler may limit this rate.

11. *Exposure Limit* Concentration of an analyte, above which worker exposure is prohibited or not recommended for a specified period of time during the workday.

    For any given analyte, there may be a number of different exposure limits based on rulings or recommendations from agencies such as the NIOSH Recommended Exposure Limit (REL),[9] the OSHA [29 CFR 1910.1000[10]] Permissible Exposure Limit (PEL), the American Conference of Governmental Industrial Hygienists (ACGIH) Threshold Limit Values,[11] Mine Safety and Health Administration (MSHA) PEL [30 CFR[10]], etc. These limits also may be international in scope and usually are expressed in the following terms:

    a. *Time-Weighted Average (TWA)* Time weighted average concentration measured over a defined time period (e.g., 15 min to 8-10 h).
       (1) *Short-Term Exposure Limit* (STEL) Time weighted average concentration measured over a limited sampling period (usually 15 min unless otherwise noted).
       (2) *Ceiling Limit* (C) Concentration which is not to be exceeded over any time period (e.g., instantaneous to ca. 5 min).

12. *Sample Stability* The ability to retrieve the analyte from the sampler after storage for a period of time under a defined set of environmental conditions.

II. METHOD DEVELOPMENT

In the development of a sampling and analytical method, there is a logical progression of events that cover the gathering of information and the preliminary experimentation for selection of analysis technique and sampling medium. Important aspects of these events are discussed below:

A. Analyte Identification - To initiate the development of a method, the identity of the analyte must be as fully defined as possible. Physical and chemical properties of the analyte should be defined so that procedures for proper handling and use of the analyte can be prepared. These also aid in establishment of analyte purity. Potential sources of this information include chemical reference books, health hazard evaluation reports, bulk sample analyses, material safety data sheets, chemical process information, etc.

B. Literature Search - Before initiating any developmental work on a method for an analyte, a thorough search of the pertinent literature should be performed. The search should reveal any other methods which have been developed for the analyte of interest or any related compounds. This search also should produce information on the toxicological properties of the analyte and any related potential health effects. This information will assist the analyst in the proper handling of samples and standards during laboratory work.

1. *Measurement and evaluation concentration ranges* to be based on several factors:

   a. Exposure limits - As stated under Section I.B.6., the evaluation range should cover, at a minimum, the range from 0.1 to 2.0 times the Exposure Limit. Higher multiples of the upper exposure limit can be added if needed (e.g., 10 times the Exposure Limit). In situations where multiple exposure limits (i.e., from different authorities) exist for an analyte, the lowest exposure limit should be used to set the lower limit of the evaluation range (0.1 times lowest exposure limit) and the highest limit used to calculate the upper limit of evaluation range (2 times the highest exposure limit). Intermediate evaluation levels should include levels at these exposure limits.

   b. Toxicology data - The toxicity of an analyte (e.g., suspected carcinogenicity) may indicate that a concentration lower than that calculated by the exposure limit should be included in the measurement and evaluation ranges.

   c. Previous monitoring data (if available) - Previous monitoring information with other methods may indicate that typical levels of the analyte may be below or above a concentration range based on the exposure limit. In this case, this lower or upper level may be included in the method evaluation.

2. *Physical and chemical properties* which may require special handling of analyte:

   a. Volatility - Volatile analytes should be kept in a closed container in a vented cabinet designed for this purpose. Solutions and neat material should be handled in a chemical fume hood to prevent exposure of the analyst and contamination of sampling media.

      If an analyte can exist both as an aerosol and a vapor, the sampler to be used must be able to collect both fractions. The sample preparation prior to analysis must also address the extraction of these fractions from the sampler.

   b. Thermal stability - The thermal stability of the analyte may require special sampling and analysis techniques, so that sample integrity is not compromised.

c. Reactivity - The reactivity of an analyte may require special means to stabilize the analyte after collection, prior to analysis and during handling. This includes both the chemical reactivity and/or photosensitivity of the analyte.

d. Any combination of the above.

3. *Information about the use of the analyte* in the industrial process should provide:

a. Identification of potential interferences

b. Process information on intermediate formation - This will help to identify any other potentially hazardous materials which might be formed in the process.

4. *Method availability* in the literature search results may indicate:

a. Applicability of the methodology to the matrix
(1) Related analytes
(2) Related methods
(3) Method for the specified analyte in a different matrix

b. Equipment needed for a specific method

C. Experimental - A sampling and analytical method can be viewed as having two related parts, namely a sampling procedure and an analysis procedure.[12] The sampling procedure includes the collection of the analyte from a known volume of air and the stabilization of the analyte on the sampling medium. The analysis procedure then deals with the subsequent recovery of the analyte from the medium and the determination of the analyte as it is removed from the sampling medium.

The success of a sampling and analytical method depends on the compatibility or interfacing of these two parts. As technology progresses, improvements can be incorporated into either part of the method to improve performance, provided that the interface remains compatible. For example, if the analysis procedure requires that the analyte be in solution in a specific organic solvent, then the collection medium for the sample can be changed (e.g., for demonstrated improvement in capacity) to other types of medium as long as the analyte is recovered at an acceptable level and ends up in the appropriate solvent. Likewise, a better measurement technique (e.g., demonstrated improvement in sensitivity or specificity) can be incorporated into an analytical method as long as compatibility with the sampler and sample workup procedures is maintained.

Since innovation is a key element in the sampling and analysis method development process, detailed experiments for the initial development of the sampling approach and optimization of the analytical procedure are better left to the discretion of the researcher. During development, it should be recognized that appropriate, statistically-designed experiments will optimize the amount of information obtained. Therefore, consultation with a statistician will be of value during this phase of the research.

The points listed below are elements that should be addressed during these method development experiments. Specific experiments for the evaluation of the sampling and analytical method are also included in Section III. In the evaluation of the method, environmental and sampling conditions which have the greatest impact on method performance should be defined and used for further experiments in the evaluation of the method.

1. *Analysis Procedure* - One of the initial decisions which must be made when developing a method is the most appropriate analytical technique for analyte determination. Consideration should be given to analytical equipment availability and associated advantages and disadvantages, operator

expertise, sensitivity requirements, analyte reactivity, interferences, reference material, etc. An outline of factors for consideration is listed below:

a. Analytical techniques for consideration may include:
   (1) Chromatography
      (a) Gas
      (b) Liquid
      (c) Thin-layer
      (d) Supercritical fluid
      (e) Ion
      (f) Electrophoresis
   (2) Spectrometry/Spectroscopy
      (a) Ultraviolet/Visible
      (b) Infrared
      (c) Fluorescence/phosphorescence/emission
      (d) Atomic
      (e) Mass
   (3) Other analysis techniques
      (a) Electrochemical
         i) Ion selective electrodes
         ii) Polarography
         iii) Other
      (b) Flow injection analysis
      (c) Titrimetric
      (d) X-Ray diffraction
      (e) Gravimetric
      (f) Microscopy
      (g) Spot test
   (4) Hyphenated combinations of above techniques
   (5) Bioanalytical
      (a) Enzyme linked immunosorbent assay
      (b) Polymerase chain reaction

b. Calibration and standardization considerations include:
   (1) Purity and availability of standards
   (2) Accuracy of standards
   (3) Preparation of standards
      (a) Instrumental detection limit
      (b) Instrumental quantitation limit
      (c) Use and availability of internal standards
      (d) Stability of standards
      (e) Linear dynamic range of the analysis technique
         i) Multiple standards, replicate analyses, same conditions as samples
         ii) Precision of replicate analyses of standards

c. Interferences in the analysis - The conditions under which the sample is collected may help to identify potential interferences.

d. Stability of the analyte:
   (1) During analysis
   (2) Prior to analysis

2. *Sampling Procedure* - The selection of sampling medium and procedure is another decision which usually is made early in the method development process. The physical state of the analyte (i.e., gas, aerosol, vapor, or combination thereof) plays an important factor on the selection of the appropriate sampler. Analytes which can exist in more than one physical state may require a combination of sampling media in one sampler for efficient collection.[13]

   Where possible, commonly-available and easily-used samplers should be investigated initially. As the preliminary testing of a sampling method progresses, further modification in the sampling medium or sampler design may be required and may impact on the analytical procedure. Again, use of commercially-available samplers where possible helps simplify sampler selection.

   a. Sampler design and media selection considerations:
      (1) Sampler design - Since industrial hygiene analytical methods are geared toward measuring personal exposure, the size, weight and convenience of the sampler are important elements in sampler design. The personal sampler should allow freedom of movement, should be unbreakable, and not prone to leakage.
         (a) Filters and holders
            i) Inlet design - Designs available include both closed and open face cassettes and inlets for inhalable particles.[14]
            ii) Filter size and type - These selections may depend on the levels of analyte expected and the compatibility of the analyte with the filter material, or the filter type with the analytical technique.
            iii) Filter holder type - This selection is based on the decisions made in i) and ii) above.
         (b) Sorbents
            i) Tubes - The amount of sorbent contained in the tube affects the capacity of the sampler for the analyte. However, as sorbent bed size increases to provide additional capacity, tube size increases and the sampler may become more cumbersome. For solvent desorption of larger sorbent beds, more solvent is required, with little or no resultant improvement in method sensitivity.
            ii) Impingers and Bubblers - These may provide only limited collection efficiency for particular analytes (e.g. aerosols of small particle size). These devices may leak and also are cumbersome, breakable and should be used only when other sampling approaches have not been successful. Potentially toxic reagents should be avoided unless they can be used safely. Reagents used should not pose any exposure hazard to the worker wearing the sampler or industrial hygienist taking the samples.
         (c) Other
            i) Reagent-coated filters or sorbents - Potentially toxic reagents should be avoided unless they can be used safely. Reagents used should not pose any exposure hazard to the worker wearing the sampler or industrial hygienist taking the samples.
            ii) Combination of different media - The combination of different types of sampling media, such as a filter and a sorbent, may be required in some instances to allow adequate collection of an analyte from air.
            iii) Passive monitors - This type of sampler requires no sampling pump but must be properly evaluated to ensure the validity of results. Passive monitor parameters and procedures for their evaluation are covered in detail in the literature.[15,16,17,18] These parameters include:
               a) Reverse diffusion - Loss of analyte from the monitor over time.
               b) Temperature - Temperature effects on the diffusion constant.

c) Shelf life - Storage of the monitor prior to use.
d) Uptake rate - The rate at which the analyte is collected by the monitor over time.
e) Parameters evaluated in a fractional factorial experiment: capacity, precision, accuracy, interferences, recovery and sample stability.

iv) Direct-reading instruments - This type of device combines sampling and measurement of the analyte into one operation and offers real time or near real time exposure monitoring information.
v) Field analyzable samplers - This approach offers the user the ability to determine the analyte at the site of sampling.

(d) Department of Transportation Regulations and Restrictions - Since samplers may be shipped to the laboratory for analysis in many instances, sampler design and media selection should consider any restrictions which might be required for commercial transport of the samplers or sampling equipment (e.g., impinger solutions; compressed gases for direct reading instruments).

(2) Artifacts on the medium and interferences in the efficient collection of the analyte on the medium - A sampling medium which contributes interferences will add complexity to the method. The media should be characterized as to what materials may be present. If potential interferences are present, a clean-up procedure for the media may be required.

III. METHOD EVALUATION

After the initial development experiments for the method have been completed and a method has been proposed, the sampling and analysis approach should be evaluated to ensure that the data collected provides accurate results. An experimental approach for this is summarized in Appendix 4 and described below in detail:

A. **Recovery of the analyte from the medium** - After preliminary experiments have identified a potential sampling medium and analytical workup, the ability to recover the analyte from the medium should be determined. A suggested experiment to accomplishment this is described below:

1. *Experimental*: Fortify sets of six[a] samplers with amounts of analyte equivalent to sampling concentrations of 0.1, 0.5, 1.0 and 2.0 (or higher) times the exposure limit for a minimum of 4 h at 0.01 to 0.20 L/min for sorbent-based samplers and 1 to 4 L/min for filter-based samplers. In some situations, other sampling rates may be more appropriate to use. At extremely low flow rates (ca. 5 mL/min), the effect of diffusion of the analyte into the sampler must be considered. Flow rates should be kept at a high enough rate to prevent diffusion from having a positive bias in the sampler.[19]

   If the analyte has a ceiling or short term exposure limit, the amount of analyte fortified should be adjusted for the shorter sampling time required for this type of exposure limit. If the sampler has a backup section, then a like number of separate backup sections should be fortified with amounts of analyte equivalent to 25% of the amount fortified on the front sections of the samplers. Samples (and backup sections) should be prepared for analysis and analyzed according to the proposed method. Results of these analyses should be expressed in terms of estimated percent recovery according to the following formula:

$$\text{Percent Recovery}_{(est.)} = \frac{(\text{Amount of analyte found on sampler})}{(\text{Amount of analyte fortified on sampler})} \times 100\%$$

   After initial analyses of the samples, the samples should be resealed and analyzed on the following day, if possible. If the sample workup procedure results in a solution of the sample, these solutions should be recapped after the initial analysis if possible and reanalyzed on the following day using fresh standards.

2. *Criterion*: Recovery should be calculated for the primary and backup media in the sampler. Although complete recovery of the analyte from the sampler is most desirable, at a minimum the estimated recovery of the analyte from the primary collection medium should be greater than or equal to 75% for levels equivalent to sampling 0.1, 0.5, 1.0 and 2.0 times the exposure limit. If recovery varies with analyte loading, results should be graphed as recovery versus loading during calibration of the method, so that appropriate correction can be made to sample results, as long as recovery is greater than 75%.[20]

   If estimated recovery does not exceed 75% at the ceiling limit, the method is not suitable for monitoring at this limit.

   Estimated recovery from any backup medium should be noted so that appropriate corrections can be applied if breakthrough of the sampler has occurred during sampling. The recovery of the analyte from the medium in the backup section of a sampler may be different from that of the front section, since the backup section of a sorbent-based sampler usually contains only half of

---

[a]See references 2, 3 and 4 for the justification of this sample size choice.

the sorbent of the primary section. If the same volume of desorption solvent is used for both the primary and backup sections of the sampler, the desorption equilibrium can be shifted, since the backup section is being desorbed by twice the volume (i.e. on a mL solvent/mg sorbent basis).[21]

Reanalysis of the samples on the day after initial analysis indicates if immediate analyses after sample preparation is required. Often when processing a large number of samples, it may be necessary to prepare the samples for analysis as a batch. In these instances, the last samples may not be analyzed for up to 24 h or more after preparation due to the time required for analysis. If samples prepared for analysis exhibit time dependent stability after desorption, analyses must be conducted within acceptable time constraints. Analysis and reanalysis results should agree within 5% of each other. This can be determined using either Student's t-test[22] or analysis of variance.[23]

B. **Stability of analyte on the sampling medium** - An experiment similar to III.A. may be performed to evaluate the stability of the analyte on the medium, with time. An additional set of samples at each of the 4 levels should be prepared and analyzed after 7 days storage at room temperature. Recovery should be similar to results above within experimental error. Discrepancies larger than those expected by experimental error indicate sample stability problems which will need correcting with additional developmental effort (e.g., refrigerated storage). Comparison of data results can be performed with statistical tests, such as the Student's t-test[22] or an analysis of variance test[23] of the means of the Day 1 and Day 7 storage results.

C. **Stability of the sampling medium prior to use** - The sampling medium should be studied to evaluate the formation of artifacts or collection of interferences or loss of collection efficiency during storage prior to use. Analysis of blank samplers over a period of time (i.e., 6 months) can give an indication of whether the sampling medium contributes artifacts or interferences to the analysis. The collection of samples from a generated atmosphere with unused, stored samplers can indicate if there is any loss of collection efficiency. Analysis results of samples collected with new and aged samplers should agree within experimental error.

D. **Sample Generation:**

1. *Feasibility of analyte generation* - As part of the development of a method, the sampling of a generated atmosphere is needed to more adequately evaluate the performance of a method.[24,25,26] If possible, the generated atmosphere should be representative of the environment encountered when sampling for the analyte in the workplace. In some instances, generation of the analyte can be quite difficult. When attempting to generate a concentration of an analyte, some areas of consideration are listed below.

2. *Impact of environmental conditions on sampler performance* and/or generation:

   a. Temperature - The effect of elevated temperature on the collection medium of a sampler may decrease the capacity of the sampler. Elevated temperature may also decompose the analyte during generation and sampling.

   b. Pressure - Reduced pressure may also reduce the capacity of a sampler.

   c. Relative humidity - High relative humidity in many instances has been observed to reduce sampler capacity.[12] In other instances it has increased sampler capacity.[27]

3. *Interference generation* - A typical interference(s) should be generated along with the analyte to approximate a typical workplace sampling environment.

4. *Particulate material generation* - Generation of particulate material can be extremely complex,[28,29] especially if particles of a particular size must be generated for the evaluation of a specified sampler inlet design. The aerodynamic performance of the generator is a factor in the generation of this type of atmosphere and should be evaluated carefully. Appropriate independent methods should be available to verify particle size if this is a critical element in the generation.

5. *Independent concentration verification* - The concentration of the generated atmosphere should be verified by replicate samples (if possible) of an independent method at each concentration level used for method evaluation or other appropriate means. A statistician should be consulted for advice on the design and sample sizes to accomplish this validation. Ideally, the independent method should not be biased and should provide an accurate estimate of the concentration generated, assuming error is randomly distributed around the mean. Also the precision and bias of the independent method should be homogeneous over the concentrations investigated.

    a. Replicates of independent method for estimation of bias - Replicate samples of the independent method should be taken to estimate accurately the true concentration of the generated atmosphere, so that any bias present in the method under evaluation can be detected. Replicate measurements at each concentration should be made to provide an estimate of the error in the measurement, which can be used in estimated bias calculations (Appendix 1 and Note 2).

    In instances where the concentration of the generator can be based only on calculations using flow rates in the generator and the amount of analyte injected, the generation system should be well characterized so that any analyte losses are known.

    b. Multiple generated concentrations to be used in method evaluation should include 0.1, 0.5, 1 and 2 times the exposure limit for the specific analyte.[12]

    c. In some instances, generation of an analyte may be extremely difficult and even hazardous. As an alternative to direct generation in these cases, generation may be simulated by fortifying samplers with an amount of analyte expected to be sampled over a specified period of time at a specific flow rate. When this is necessary, fortification of the sampler by vaporization of a known amount of analyte onto the sampling medium is an appropriate method, since this approach more closely approximates a generated atmosphere. The alternative of direct application of a solution of analyte onto the collection medium is less desirable but may be necessary in some instances.

    After fortification, air conditioned at both high and low humidity should be drawn through samplers at the flow rate and time period used in the calculations for the amount of analyte expected to be collected. In the method report, the fact that samples were not collected from a generated atmosphere should be discussed.

E. Capacity of the sampler and sampling rate - To determine the applicability of the sampling method, the capacity of the sampler should be determined as a function of flow rate and sampling time. This is particularly important if the analyte has both a short-term limit and a time weighted average. A suggested experiment to help determine the capacity of the sampler is given below:

1. *Experimental*: Sampling rates typical for the media selected should be used. Typically, these may range from 0.01 - 0.20 L/min for sorbent tube samplers to 1-4 L/min for 37-mm filter cassette samplers. Other types of samplers may require different flow ranges. Sampling should be performed at three different flow rates covering the range discussed above for the particular sampler type, unless the sampler is designed to operate at only one flow rate. Sampling times should range from 22.5 min for short term exposure limits to 900 min (15 h) for time weighted

averages. Shorter sampling times (e.g., 7.5 to 22.5 min) may be used for ceiling measurements. Flow rates should be based on accurately calibrated sampling pumps or critical orifices. The amount of analyte collected at the lowest flow rate and shortest sampling time should be greater than the LOQ. The generated concentration used for capacity determination should be at least 2 times the highest published exposure limit and verified by an independent method. Sampling should be conducted at ambient, elevated (>35 °C) and low (<20 °C) temperatures to assess the effect of temperature on sampling. To assess the effect of humidity on capacity, sampling should be performed at both low and high humidities ($\leq$ 20% and $\geq$ 80%, since both have been observed to affect capacity.[27,12] Triplicate samplers at three different flow rates should be included to verify capacity at different humidities. For samplers which contain backup sampling media, only the front section of the sampler should be used. A means is required to quantitate analyte in the effluent from the sampler. This may involve the use of a backup sampler, continuous monitor or other appropriate means which can provide a measure of analyte concentration in the sampler effluent (ca. 1 - 5% of the influent concentration). If the mass of analyte found on the backup sampler totaled 5% of the mass found on the front sampler or the effluent concentration of the sampler contained 5% of the influent concentration, breakthrough has occurred and the capacity of the sampler has been exceeded.

If the analyte is a particulate material and collected with a filter, the capacity of the filter is defined by the pressure drop across the sampler or by the loading of the filter.

2. *Criterion*: If the collection process is based primarily on adsorption, breakthrough time should be proportional to the inverse of the flow rate.[30] This relationship can be checked by plotting the 5% breakthrough time versus the inverse of the flow rate. If the resulting plot is a straight line, then this relationship should hold for all flow rates in the flow rate range studied. Some nonlinearity in the plot may be noted due to experimental variability and assumptions made to simplify the relationship of breakthrough time and flow rate. Results from these experimental trials should provide a prediction of the capacity of the sampler at various flow rates and sampling times. If the flow rates and sampling times used in the experiment do not provide for sufficient capacity, a lower flow rate range may have to be studied and the experiment repeated.

   With samplers which use reagents for collection of the analyte, the amount of the reagent in the sampler will also be a limiting factor in the capacity of the sampler, based on the stoichiometry of the reaction. Other factors, such as residence time in the sampler and kinetics of reaction between analyte and reagent, may affect the capacity of this type of sampler. For filter-based samplers, pressure drop should be less than 40 inches (1016 mm) of water for total loading less than 2 mg.

   The combined temperature and humidity conditions which reduces sampler capacity to the greatest extent should be used in all further experiments. The Maximum Recommended Sampling Time (MRST) for a specific flow rate is defined as the time at which sampler capacity was reached multiplied by 0.667. This adds a measure of safety to this determination. The relationship of breakthrough time with flow rate can be used to adjust flow rates to allow specific sampling times.

F. **Sampling and Analysis Evaluation** - To assess the performance of a method, certain additional experimental parameters should be evaluated through a series of defined experiments. For each method under consideration, the goals are to estimate method accuracy and to determine if the method meets the 25% accuracy criterion (i.e., either to determine that there is 95% confidence that the method has an accuracy of no more than 25% of the true amount measured with probability of 0.95, or to determine that there is 95% confidence that the method has an accuracy of less than 25%). With the recommended sample sizes, this goal is realistic for methods with accuracies of 12.5% or less or 40% or more. For methods with accuracies from 12.5% to 40%, the results are likely to be

inconclusive unless larger sample sizes are used. The closer the method accuracy is to 25% the larger the sample sizes must be to have great chance to determine whether the method does or does not meet the 25% accuracy criterion. A statistician should be consulted for planning the experiment if the method accuracy is thought to be greater than 12.5% or if a conclusive result is desired even for methods with accuracies close to 25%. Experiments directed at achieving this goal should include evaluation of the following:

1. *Environmental parameters*

   a. Experimental: The effect of environmental conditions on sampling efficiency of the medium listed below can be evaluated by a multi-level factorial design. The relative humidity, flow rate, and sampling times determined in the experiment described above to most severely limit sampler capacity should be used in these experimental runs. At a minimum, the effect of concentration on method performance should be investigated. Three sets of twelve[a,b] samples should be collected from an atmosphere containing concentrations of 0.1, 1.0 and 2.0 times the exposure limit at the humidity determined above to reduce sampler capacity for the MRST determined in the preceding experiment (III.E.). If the analyte has a short-term or ceiling exposure limit in addition to a 8-hour time weighted average, an additional 12[c] samplers should be collected at the STEL or C limit for the recommended sampling period at the appropriate flow rate. Environmental factors that might be evaluated at multiple levels are listed below:
      (1) Interference - Potential interferences in the work environment should be included in the generation experiments to assess their impact on method performance. Concentrations up to 2 times the exposure limit value for the interference should be included.
      (2) Other environmental factors may be studied, but will require a more comprehensive experimental design.

   b. Application: The effect of environmental parameters on method performance should be assessed. The factorial design used to evaluate these factors should define which exert a significant effect on analyte recovery (See Appendix 5 for an example of this experimental design and data analysis). Those factors which are found to influence analyte recovery should be investigated further to determine if their impact is predictable. If so, corrections can be made to the data using the predictive equations which would result from these further investigations. If these effects are not predictable, the utility of the method will be limited based on the conditions defined by this experiment.

      If only concentration is evaluated, the analyte recovery should be the same at all levels after correctable biases have been included, such as estimated recovery.

---

[a] Sets of nine (9) can suffice if the hyperbolic approximation is used (subject to the considerations mentioned in footnote b). See Appendix 1 for a description of the alternatives.

[b] This is a guide for "typical" methods with precisions (including an allowance for a 5% pump factor) between 7% and 9%. If the available data indicates that the precision of the method under evaluation falls outside this range, a statistician can calculate expected Type II (false non-acceptance) probabilities for samples of twelve and/or more or less per concentration. This information can be used to plan the sample sizes to be used.

[c] Nine (9) can suffice if the hyperbolic approximation is used. See Appendix 1 for a description of the alternatives.

G. **Pressure drop across sampler** - The pressure drop across the sampler should not be so great as to limit sample collection times to $\leq$ 10 h. For analytes with only an STEL value, this $\leq$ 10 h recommendations can be reduced to $\leq$ 1 h.

H. **Sample Stability** - To assess sample stability, samples should be collected from a generated atmosphere, stored under defined conditions, and analyzed at specified time periods. A suggested experiment is described below:

1. *Experimental*: A concentration of 0.5 times the lowest exposure limit should be sampled with 30[a] samplers for a minimum of 1/2 the MRST. The humidity and temperature of the generator should be at the same level as defined in III.E. to reduce sample capacity. The samplers should be divided randomly into one group of 12[b], one group of 6, and four groups of 3, with the group of 12[a] analyzed as soon after collection as possible (Day 0). The group of 6 should be analyzed after 7 days. The four remaining sets of 3 samples should be analyzed after 10, 14, 21, and 30 days. The conditions of storage are determined by the nature of the analyte. If there is an indication of analyte instability on the sampling medium, refrigeration of the samplers may be required. However, storage for the first seven days should be at room temperature.

2. *Criterion*: If the average analysis results of the set of samplers analyzed on day 7 differs from the set analyzed on day 0 by more than 10%, the method does not meet the sample stability criterion. Samples should be stable for a minimum of 7 days under ambient conditions to allow samples to be shipped to a laboratory for analysis.

   Either additional precautions, such as shipment on ice and refrigerator storage, may be required or the method may have to be modified to address this problem. If a plot of recovery versus time indicates that recovery decreased by more than 10% after the initial 7-day storage period, sample instability is a problem. If samples need to be stored for longer periods, more restrictive storage conditions are required. Remedial action, such as cold storage may solve this longer term storage problem. After remedial precautions have been instituted in the method, the sample stability of the method must be redetermined.

I. **Precision, Bias, and Accuracy** - Results from the analyte recovery experiment and the sampling and analysis experiments (e.g., the environmental parameters experiments) and the sample stability experiment, can be used for the estimation of precision, bias, and accuracy of the method. To further define the accuracy, Appendix 6 contains an algorithm for the calculation of an estimate of method accuracy using PC-SAS programming language.[31] This can also be accomplished using the nomogram in Figure 1. A DOS program, ABCV.EXE is also available to calculate any of the three attributes -- accuracy, bias, precision -- as a function of the other two[c].

1. *Criterion*: Sampler results from the environmental parameters experiment at the 0.1, 1.0 and 2.0 times the exposure limit value (III.F.), the sampler stability experiment (III.H) (at 0.5 times the exposure limit) and the analytical recovery experiment (III.A.) are used in the calculations of method precision. The calculations for $\hat{S}_{rT}$,i.e., the estimate of S_{rT}, are described in Appendix 2.

---

[a] This number must be adjusted to the sample sizes chosen at each concentration for the accuracy evaluation. This assumes 12 per concentration. If 9 per concentration is used, then 27 will suffice. If the sample size for the accuracy evaluation is n, then this number should be n + 6 + (4 x 3).

[b] See the footnote on the previous page.

[c] The program, ABCV.EXE, is available from the Division of Physical Sciences and Engineering of the National Institute for Occupational Safety and Health.

Before obtaining a pooled estimate of S_{rT} from the four sets of samplers, the homogeneity of the precision over the range of concentrations studied should be checked using a test, such as Bartlett's test.[2,3,4] If the precision is not found to be constant over concentrations, the sample set collected at 0.1 x exposure limit should be removed and Bartlett's test recalculated. Homogeneity of the method precision is an assumption required not only to obtain a pooled estimate of S_{rT}, but this assumption is also required to construct confidence interval estimates of the bias applicable to the entire range of all concentration levels studied.

Bias is assumed to be homogeneous over the evaluation range. This assumption should be tested by estimating the bias at each concentration level and testing these for homogeneity using the procedures describe in Appendix 1. Method bias should be less than 10% and a test for this is described in Appendix 1. The best single and 95% confidence interval estimates of the method bias are computed using procedures in Appendix 1. The estimate of the worst-case bias is the upper limit of the 95% confidence interval estimate as defined in Appendix 1 when the point estimate is positive (It is the lower 95% limit when the point estimate is negative). The estimate of the best-case bias of the method is the lower limit of the 95% confidence interval estimate as defined in Appendix 1 when the point estimate is positive (It is the upper 95% limit when the point estimate is negative) but only if the confidence interval does not include zero; otherwise it is 0.0.

If the individual S_{rT}s are not found to be unequal, then their estimates can be pooled together to provide both the best single estimate, ($\hat{S}_{rT}$), and the 95% confidence interval limit estimates of the precision of the method as described in Appendix 1.

The bias and precision estimates are then used as described in Appendix 1 to compute the best single and 90% confidence interval estimates of the accuracy. The latter is bounded by the 5% and 95% confidence limit estimates of the accuracy. If both of these estimates are less than 25%, then there is 95%[a] confidence that the method accuracy is no larger than 25% and the method is declared as fulfilling the 25% accuracy criterion. If both of these estimates are greater than 25%, then there is 95%[a] confidence that the method accuracy is greater than 25% and the method is declared as not fulfilling the 25% accuracy criterion. However, if the 90% confidence interval estimate for method accuracy includes the value of 25%, the results are inconclusive.

If the results for 4 concentration levels are inconclusive or the method fails the 25% accuracy criterion, then the set of samples collected at 0.1 x exposure limit should be excluded from the data set. The pooled $\hat{S}_{rT}$ and the bias should be recalculated on this reduced data set before performing the accuracy analysis described above.

For the 12 samplers[b] collected at the ceiling limit, the accuracy analysis described above should be repeated using only the data collected at the ceiling limit.

J. Field Evaluation - While field evaluation is not required in method evaluation, it does provide further test of the method. Conditions which exist in the field are difficult to reproduce in the laboratory. Also unknown variables may affect sampling results when field samples are taken.

1. *Experimental* - Both the collection of area samples and personal samples should be included in the field evaluation of the method. Area samples should provide an estimate of field precision and bias. Personal samples may confirm these values and also provide a means to assess the

---

[a] See the footnote in Appendix 1 in Section V.A. for a discussion of the confidence coefficients.

[b] See the discussion and footnotes above concerning sample sizes per concentration.

utility of the method. A statistical study design should be prepared based on the variability of the method and the statistical power required to observed differences between the independent method and the method under evaluation.[32]

If a power study is not feasible, a minimum of 20 pairs of samplers from the method under study and an independent method should be used for personal sampling. However, it is highly recommended that a power study be done to determine appropriate sample sizes. Placement of the samplers on the workers should be random to prevent the biasing of results due to the "handedness" of the worker. Workers sampled should be in areas where both low and high concentrations of the analyte may be present.

Sets of a minimum of 6 area samplers paired with independent methods should be placed in areas of low, intermediate and high analyte concentration. If the atmosphere sampled is not homogeneous, precautions may have to be taken to ensure that all samplers are exposed to the same concentrations.

2. *Analysis* - Field precision and bias of the area sampler results of the method under study should compare with laboratory evaluation results, provided that precautions have been taken to ensure that all samplers have been exposed to the same homogeneous atmosphere. This can be done by using field exposure chambers, such as those described in the literature.[33,34] However, the use of field exposure chambers for aerosol contaminants may alter the physical characteristics of the aerosol. In these cases, the field exposure chamber may not be appropriate. Differences in precision and bias can be investigated using either Student's t-test[22] or analysis of variance.[23] Sources of variation should be studied and corrections implemented where necessary. Evaluation of personal sampler results should be done cautiously, since observable differences may be due to work practices or other situations which are beyond the control of the method.

IV. DOCUMENTATION

A. Final report - Development and evaluation research on an analytical method should be documented in a final report. This report can be in two forms:

1. *Technical report* (acceptable method development) - This report documents the successful development of the analytical method. This report may be prepared in a format appropriate for submission to a peer-reviewed journal for publication. Appendix 7 provides an example of a back-up data report for a method in the NIOSH Manual of Analytical Methods (NMAM).

2. *Failure report* (no acceptable method developed) - This report documents the research performed on an attempted method development for an analyte or analytes. The report should describe the failure of the method as well as other areas of the method research that were successful. Recommendations to solve the failure of the method may be included. Appendix 8 provides an example of a failure report for a method studied under the Standards Completion Program.

B. Analytical method prepared in appropriate format - The format of the resulting analytical method should provide clear instructions for the use of the method. Sampling, sample workup, and analysis procedures should be clearly described. The necessary equipment and supplies for the method should be listed clearly in the method. A summary of the evaluation of the method should be included, as well as a discussion of method applicability and lists of interferences and related references. A method in the format of the Fourth Edition of NMAM is included in Appendix 9.

REFERENCES

1. Occupational Safety and Health Act of 1970 (P.L. 91-596): Title 84 U.S. Code, Section 6.

2. Anderson CC, Gunderson EC, Coulson DM [1981]. Sampling and Analytical Methodology for Workplace Chemical Hazards. In: Choudhary G, ed. Chemical Hazards in the Workplace. Washington, DC: American Chemical Society, pp. 3-19.

3. Busch KA, Taylor DG [1981]. Statistical Protocol for the NIOSH Validation Tests. In: Choudhary G, ed. Chemical Hazards in the Workplace. Washington, DC: American Chemical Society, pp. 503-517.

4. Gunderson EC, Anderson CC, Smith RH, Doemeny LD [1980]. Development and Validation of Methods for Sampling and Analysis of Workplace Toxic Substances with a Statistical Appendix by KA Busch. Cincinnati, OH: U.S. Department of Health and Human Services, Public Health Service, Centers for Disease Control and Prevention, National Institute for Occupational Safety and Health, Division of Physical Sciences and Engineering. DHHS (NIOSH) Publication No. 80-133.

5. Fowler WK, Bradley Jr EL [1990]. Laboratory control of analytical method inaccuracy with respect to a given inaccuracy standard. Am Ind Hyg Assoc J 51:127-131.

6. Quality Assurance & Laboratory Operating Procedures of the Measurements Research Support Branch, Methods Research Branch, Division of Physical Sciences and Engineering, National Institute for Occupational Safety and Health (latest revision dated October 28, 1992).

7. Shotwell HP, Caporossi JC, McCollum RC, Mellor JF [1979]. A validation procedure for air sampling-analysis systems, Am Ind Hyg Assoc J 40:737-742.

8. European Committee for European Standardization (CEN), Technical Committee 137, Working Group 2 [1992]. Workplace atmospheres - General requirements for the performance of procedures for measurement of chemical agents. Document N97.

9. NIOSH [1992]. NIOSH recommendations for occupational safety and health. Compendium of policy documents and statements. Cincinnati, OH: U.S. Department of Health and Human Services, Public Health Service, Centers for Disease Control and Prevention, National Institute for Occupational Safety and Health, DHHS (NIOSH) Publication No. 92-100.

10. CFR. Code of Federal regulations. Washington, DC: U.S. Government Printing Office, Office of the Federal Register.

11. American Conference of Governmental Industrial Hygienists [1992]. 1992-1993 Threshold Limit Values for chemical substances and physical agents and biological exposure indices. Cincinnati, OH: American Conference of Governmental Industrial Hygienists.

12. Melcher RG, Langner RR, Kagel RO [1978]. Criteria for the evaluation of methods for the collection of organic pollutants in air using solid sorbents. Am Ind Hyg Assoc J 39:349-361.

13. Streicher RP, Kennedy ER, Lorberau CD [1994]. Strategies for the simultaneous collection of vapours and aerosols with emphasis on isocyanate sampling. Analyst 119:89-97.

14. Vincent, JH [1988]. Sampling for total aerosol as defined by the inspirability criterion. In Advances in Air Sampling, American Conference of Governmental Industrial Hygienists. Chelsea, MI: Lewis Publishers, Inc. pp. 17-24.

15. Cassinelli ME, Hull RD, Crable JV, Teass AW [1987]. Protocol for the Evaluation of Passive Monitors. In: Berlin A, Brown RH, Saunders KJ, eds. Diffusive Sampling - An Alternative Approach to Workplace Air Monitoring. London, England: Royal Society of Chemistry, Burlington House, pp. 190-202.

16. Bartley DL, Doemeny LJ, Taylor DG [1983]. Diffusive Monitoring of Fluctuating Concentrations. Am Ind Hyg Assoc J 44:241-247.

17. European Committee for European Standardization (CEN), Technical Committee 137, Working Group 2 [1992]. Workplace atmospheres - Requirements and test method for diffusive samplers for the determination of gases and vapours (prEN 838).

18. Bartley DL, Deye GJ, Woebkenberg ML [1987]. Diffusive monitor test: Performance under transient conditions. Appl Ind Hyg 2:119-122.

19. Heitbrink WA [1983]. Diffusion effects under low flow conditions. Am Ind Hyg Assoc J 44:453-462.

20. Saalwaechter AT, McCammon Jr CS, Roper CP, Carlberg KS [1977]. Performance of the NIOSH charcoal tube technique for the determination of air concentrations of organic vapors. Am Ind Hyg Assoc J 38:476-486.

21. Posner JC, Okenfuss JR [1981]. Desorption of organic analytes from activated carbon. Am Ind Hyg Assoc J 42:643-646.

22. Box GEP, Hunter WG, Hunter JS [1978]. Statistics for Experimenters. New York, NY: John Wiley & Sons, Inc., pp.49-53.

23. Box GEP, Hunter WG, Hunter JS [1978]. Statistics for Experimenters. New York, NY: John Wiley & Sons, Inc., pp. 170-174.

24. Woodfin WJ [1984]. Gas and Vapor Generating Systems for Laboratories. Cincinnati, OH: U.S. Department of Health and Human Services, Public Health Service, Centers for Disease Control and Prevention, National Institute for Occupational Safety and Health, Division of Physical Sciences and Engineering. DHHS (NIOSH) Publication No. 84-113.

25. Nelson GO [1971]. Controlled Test Atmospheres. Ann Arbor, MI: Ann Arbor Science Publishers.

26. Nelson GO [1992]. Gas Mixtures: Preparation and Control. Ann Arbor, MI: Lewis Publishers.

27. Cassinelli ME [1991]. Development of a solid sorbent monitoring method for chlorine and bromine in air with determination by ion chromatography. Appl Occup Environ Hyg 6:215-226.

28. Wilke K, Ed. [1980]. Generation of Aerosols and Facilities for Exposure Experiments. Ann Arbor, MI: Ann Arbor Science Publishers, Inc.

29. Hinds WC [1982]. Aerosol Technology. New York, NY: John Wiley and Sons, pp. 379-395.

30. Jonas LA, Rehrmann JA [1973]. Predictive equations in gas adsorption kinetics. Carbon 11:59-64.

31. SAS Institute, Inc. [1988]. SAS Language Guide for Personal Computers, Release 6.03 Edition, Cary, NC.

32. European Committee for European Standardization (CEN), Technical Committee 137, Working Group 2 [1993]. Workplace atmospheres - Requirements and test method for pumped sorbent tubes for the determination of gases and vapours (prEN 1076).

33. Cassinelli ME, Hull RD, Cuendet PA [1985]. Performance of sulfur dioxide passive monitors. Am Ind Hyg Assoc J 45:599-608.

34. Kennedy ER, Smith DL, Geraci, Jr CL [1985]. Field Evaluations of Sampling and Analytical Methods for Formaldehyde. In Turoski V, ed. Formaldehyde -Analytical Chemistry and Toxicology. Washington, DC: American Chemical Society, pp. 151-159.

FIGURE 1 - Nomogram Relating Accuracy to Precision and Bias

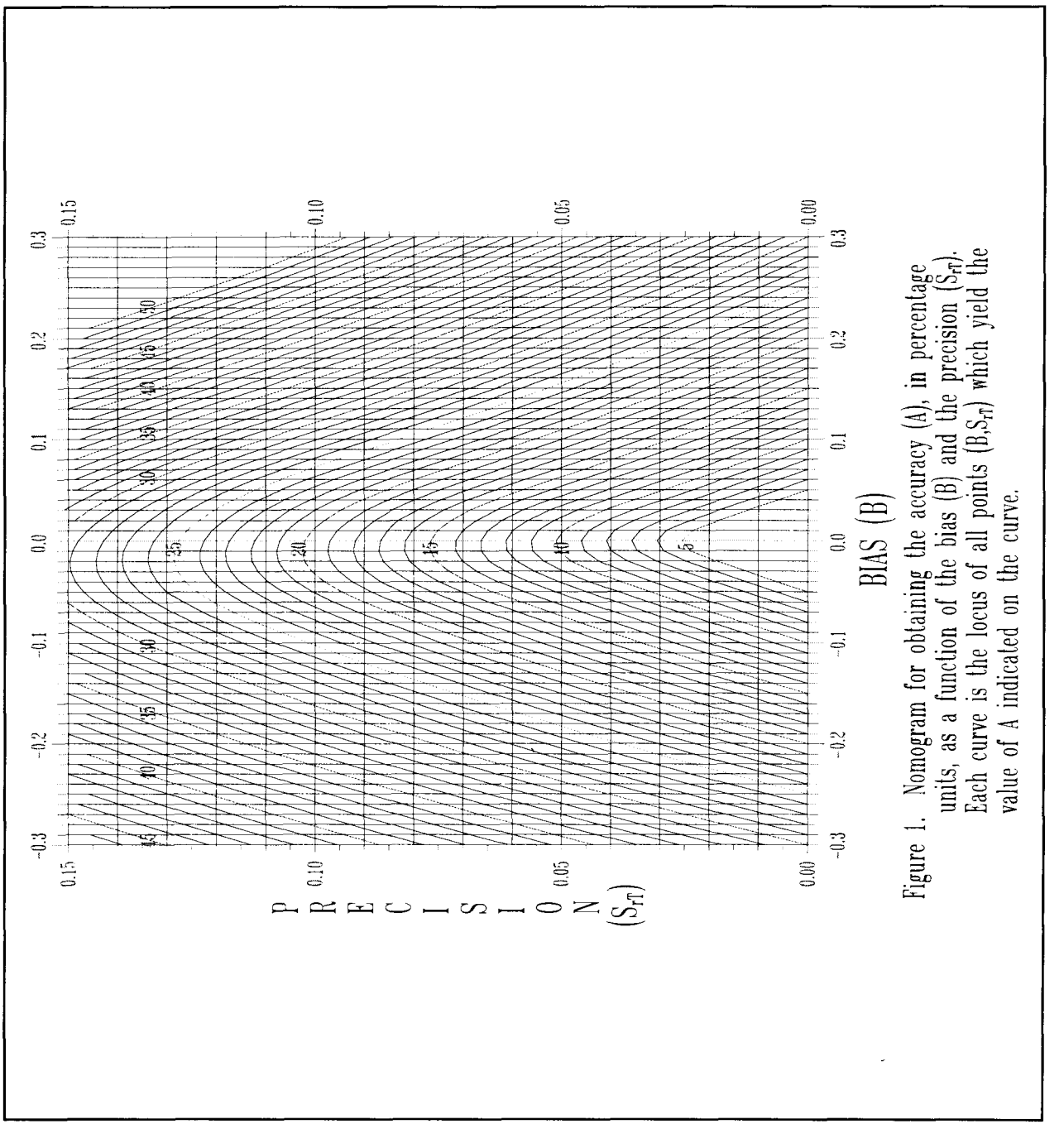


Figure 1. Nomogram for obtaining the accuracy (A), in percentage units, as a function of the bias (B) and the precision (S_{rT}). Each curve is the locus of all points (B,S_{rT}) which yield the value of A indicated on the curve.

TABLE I - Values of the Bias (B) and the Precision (S_{rT}) Required to Obtain Designated Values of Accuracy (A) in Percentage Units[a].

A (%)	B (%)	S_{rT} (%)
5	-3.5	0.9450*
5	-2.5	1.5589*
5	0.0	2.5511*
5	2.5	1.4829*
5	3.5	0.8811*
10	-7.5	1.6432*
10	-5.0	3.1999*
10	0.0	5.1022
10	5.0	2.8952*
10	7.5	1.4139*
15	-10.0	3.3777*
15	-5.0	6.3814
15	0.0	7.6530
15	5.0	5.7736
15	10.0	2.7636*
20	-10.0	6.7554
20	-5.0	9.4476
20	0.0	10.2043
20	5.0	8.5478
20	10.0	5.5271
25	-10.0	10.1284
25	-5.0	12.3869
25	0.0	12.7548
25	5.0	11.2072
25	10.0	8.2869
30§	-15.0&	10.7287
30§	-7.5	14.5544
30§	0.0	15.3061
30§	7.5	12.5236
30§	15.0&	7.9299
35§	-15.0&	14.3038
35§	-7.5	17.5897
35§	0.0	17.8574
35§	7.5	15.1353
35§	15.0&	10.5724

* Below the minimum attainable precision with a 5% pump correction.
§ Does not fulfill the Accuracy Criterion (see I.A.1).
& Does not fulfill the bias criterion (see I.A.3).

[a] Note: the values shown in this table are population or theoretical values.

TABLE II: Acronyms used in this document.

Abbreviation	Definition
AC	Accuracy criterion: A method must give a result that is within $\pm$ 25% of the true concentration with a probability of 0.95 for an individual observation
ACGIH	American Conference of Governmental Industrial Hygienists
B	Bias
C	Ceiling Limit
LOD	Limit of Detection
LOQ	Limit of Quantitation
MRST	Maximum Recommended Sampling Time
NIOSH	National Institute for Occupational Safety and Health
NMAM	NIOSH Manual of Analytical Methods
OSHA	Occupational Safety and Health Administration
PEL	OSHA Permissible Exposure Limit
RDL	Reliable Detection Limit
REL	NIOSH Recommended Exposure Limit
SAS	Statistical analysis software program
S_{r1}	Analytical relative standard deviation
S_{r2}	Sampling relative standard deviation
S_{rT}	Total relative standard deviation
S_r	Relative standard deviation
STEL	Short Term Exposure Limit
T	True concentration
TWA	Time Weighted Average
σ	The variability of population measurements about the mean
μ	The mean value of a given concentration
$\hat{S}_{rT}$	Estimated total relative standard deviation

I. INTRODUCTION

A. Types of Estimate The procedures to be described will produce estimates of the bias, precision, or accuracy. Each estimate will be the "best estimate" if one exists, or at least a reasonable one. The "best estimate," or simply the estimate, is the single value which is the best choice among all such values as the estimate of the bias, precision, or accuracy on the basis of some set of criteria. When possible, the standard error of the estimate will be estimated. Finally, in all cases a confidence interval estimate will be calculated (all confidence interval estimates will be two-sided unless explicitly described as one-sided). The bounds of these intervals are termed confidence statistics, e.g., the bounds for a 90% (symmetric) confidence interval would be a 5% confidence statistic (for the lower bound) and a 95% confidence statistic (for the upper bound). A ϕ x 100% confidence statistic is defined by the property that ϕ is the probability that such a statistic is greater than the quantity of interest, e.g, bias, precision, or accuracy. For such a statistic, there is ϕ x 100% confidence that the observed value of the statistic is greater than the quantity of interest, e.g., accuracy. On the other hand, there is (1-ϕ) x 100% confidence that it is less than the quantity of interest, e.g., accuracy.

In the case of the bias and the precision, the intervals will be estimated with 95% confidence. This will provide 2.5% and 97.5% confidence statistics for the bias and precision. A 90% confidence interval estimate of the accuracy will be obtained. This yields 5% and 95% confidence statistics for the accuracy[a]. If the 95% confidence statistic for the accuracy is 25% or less, then there is at least 95% confidence that the method satisfies the 25% accuracy criterion. If the 5% confidence statistic for the accuracy is 25% or more, then there is at least 95% confidence that the method does not fulfill the 25% accuracy criterion.

B. Source of Data for Estimation The data to be used for estimation of the bias in Section II include the measurements by the analytical method (method) under evaluation (also termed the study method) described under "Precision, Bias and Accuracy" (III.I.) and "Sample Stability" (III.H.) of the main text of these Guidelines. An observation is denoted by the symbol, C_{iMj}, which denotes the jth of n_{iM} measurements by the study method at the ith of k concentrations. These data also include either the estimated concentration based on the calculated concentration in the generation system or the measurements from an independent method used to estimate or verify the amount. An independent method measurement is denoted by the symbol, C_{ij}, which denotes the jth of n_i measurements by the independent method at the ith concentration.

Section II also describes tests of the assumption of homogeneity of bias over concentrations. Finally, Section III describes the test for the assumption that the absolute bias is less than 10%. The data used for estimation of the study method precision, S_{rT}, include the result from Appendix 2 calculations, denoted by the symbol, $\hat{S}_{rT}$, and the degrees of freedom for that statistic. Confidence statistics for the precision are described in Section IV.

It is more efficient and convenient to estimate accuracy indirectly using the estimates of and confidence statistics for the bias and for the precision. These are described in Section V.

C. Statistical Assumptions It is assumed that the $\{C_{iMj}\}$ are statistically independent and normally distributed random variables with mean, μ_{iM}, and variance, σ_{iM}, at the ith concentration given by:

---

[a] Because of the procedures used to construct the limits for the 90 percent interval estimate for the accuracy, the following statements apply to the *a priori* probability associated with each: The probability that the 95 percent limit is greater than the accuracy is at least 0.95. The probability that the 5 percent limit is less than the accuracy is at least 0.95 so the probability that the 5 percent limit is greater than the accuracy is at most 0.05.

$$(1) \quad \mu_{iM} = B_i^* \theta_i$$
$$= (B_i + 1) \theta_i$$
$$\sigma_{iM}^2 = \mu_{iM}^2 S_{ri}^2$$

where: B_i is the study method bias;

$B_i^* = B_i + 1$ is the unsigned bias;

θ_i is the concentration; and
S_{ri} is the precision (relative standard deviation).

Further, it is assumed that the $\{C_{ilj}\}$ are statistically independent and normally distributed with mean, μ_{il}, and variance, σ_{il}, at the ith concentration given by:

$$(2) \quad \mu_{iI} = \theta_i$$
$$\sigma_{iI}^2 = S_{rIi}^2 \theta_i^2$$

where S_{rIi} is the precision of the independent method.

It is assumed that the $\{S_{ri}\}$ and the $\{S_{Ii}\}$ are sufficiently small that the logarithms of the measurements (plus some positive constant -- usually 0 -- so that the logarithms exist) by both methods are approximately normally distributed[1].

Finally, for all confidence statistics pertaining to the bias for the entire range of concentrations studied, it is assumed that the study method precision, and, where applicable, the independent method precision are homogeneous, i.e., constant, over that range of concentrations.

II. ESTIMATED BIAS The estimator for the bias of a method depends on several conditions. These are:

1) whether the bias is known without error or not;
2) if not known, whether the concentration is known without error or not;
3) if not, whether the independent method used to estimate the concentration is paired with the method under evaluation when measurements are taken or not.

Estimators are defined for all these cases for any specific concentration and for a bias assumed to be constant for all k concentrations. These are described below.

A. Bias Is Known If the bias or an upper limit (this should not be confused with an upper "confidence" limit) for the bias is known without error, then that value is used throughout as the bias for the method. The bias for the ith concentration is given as

$$(3) \quad B_i = (\mu_{iM}/\theta_i) - 1$$

where: μ_{iM} = the mean of the method under evaluation, and
θ_i = the measured concentration.

(However, either or both μ_{iM} and θ_i may be unknown.) When the bias is known to be constant over several concentrations, then its value is given by Equation 2 of the main text of these Guidelines. This estimator is also a $(1-\alpha)$ x 100% confidence statistic for the bias where $\alpha = 0$, e.g., the confidence is 100%. Thus, a 100% confidence interval "estimate" is the known bias.

Remark: If the bias is estimated but it is used as if it were known, any confidence statement about accuracy may be invalid and incorrect. Thus, there may not be 95% confidence that a method fulfills the 25% accuracy criterion if this conclusion requires that the bias be known when it is estimated. It is not enough to find that the bias estimate is "nonsignificant" or its difference from zero is not statistically significant at the 5% level. That result indicates that there is not 95% confidence of a non-zero bias. The issue is that the assumption of a known bias is contradicted by the fact that an estimate with a standard error is used. Moreover,a second issue is that the non-significance of the test does not establish confidence of a bias of 0% unless there is good prior reason for predicting that the bias should be 0% (or unless the power of the test is sufficiently great so that the non-significant result indicates a negligible bias). If experimental data are used to determine a value for the bias, that value should be treated as estimated for valid confidence statements about the accuracy.

B. Bias Is Not Known There are three cases when the bias is not known:

1) the measured concentration is known so that the uncertainty in the bias is entirely a function of the uncertainty in the mean of the study method;
2) when both the study method mean and the measured concentration are not known; and
3) the bias is estimated with data that are independent of the data used to estimate the precision, i.e., a case not treated in these Guidelines.

1. *Measured Concentrations Known Without Error* Two sets of estimates are presented. For the first set, it is simply assumed that the concentrations, $\{\theta_i\}$, for the k concentrations are known without specifying how. It is the user's responsibility to ensure that the estimators given are statistically justified and scientifically valid. It is not enough that the concentrations are planned or targeted (unless the experimental setup described below or an equivalent is used): the realized concentrations must be known to be identical with the planned ones. A "nonsignificant" difference between the planned concentration and the mean of an independent method is not sufficient. Such a result establishes that there is not 95% confidence that a difference exists: the issue is high confidence of no difference. Next, a second set of estimators is presented along with an experimental setup and a statistical model that would justify the assertion that the concentrations to be measured are known. This experimental setup and the statistical justification for the estimators are fully described in the second note at the end of this appendix.

a. Concentrations Assumed Known

(1) Estimates When the concentration to be measured is known without error, the bias estimate for the ith concentration is given by:

$$(4)\quad \hat{B}_i = \frac{\bar{C}_{iM.}}{\theta_i} - 1, \quad \textit{for } i = 1,2,\ldots,k \textit{ concentrations},$$

$$\textit{where} \quad \bar{C}_{iM.} = \frac{\sum_{j=1}^{n_{iM}} C_{iMj}}{n_{iM}}$$

$$= \textit{the average at the ith measured concentration.}$$

C_{iMj} = the jth of the n_{iM} measurements on the method under evaluation at the ith concentration;

θ_i = the ith concentration (this is a known constant);

n_{iM} = the number of measurements on the method under evaluation at the ith concentration;

Or, in terms of the measurements:

$$(4a) \quad \hat{B}_i = \frac{\sum_{j=1}^{n_{iM}} C_{iMj}}{n_{iM}\,\theta_i} - 1$$

The estimate of the standard deviation of the measurements, termed σ_{iM}, is given by:

$$(5) \quad \hat{\sigma}_{iM} = \sqrt{\frac{\sum_{j=1}^{n_{iM}} (C_{iMj} - \bar{C}_{iM.})^2}{n_{iM}-1}}$$

$$= \textit{The Estimated Standard Deviation of } \{C_{iMj}, j=1,\ldots,n_{iM}\}$$

$$\textit{which has } n_{iM} - 1 \textit{ degrees of freedom.}$$

The standard error of the estimated bias at concentration, i, is given by:

$$(6) \quad \hat{\sigma}_{\hat{B}_i} = \frac{\hat{\sigma}_{iM}}{\theta_i \sqrt{n_{iM}}}$$

The 95 percent confidence interval estimate for the bias for the ith concentration is given by the interval ($\hat{B}_i^{0.025}$, $\hat{B}_i^{0.975}$) , where these limits are the 0.025 and 0.975, respectively, confidence statistics for the bias, given by:

$$(7) \quad \hat{B}_i^{0.025} = \hat{B}_i + t_{df_i}^{0.025}\,\hat{\sigma}_{\hat{B}_i}$$
$$\hat{B}_i^{0.975} = \hat{B}_i + t_{df_i}^{0.975}\,\hat{\sigma}_{\hat{B}_i}$$

df_i = the degrees of freedom for the estimate of σ_{iM} , usually = n_{iM} -1

$t_{df_i}^{1-\alpha}$ = (1-α)th percentile of the t distribution with df_i degrees of freedom.

Thus, the limits for the 95 percent confidence interval estimate for the bias at the ith concentration can be written in terms of the measurements as follows:

$$(7a)\quad \hat{B}_i^{0.025} = \frac{\sum_{j=1}^{n_{iM}} C_{iMj}}{\theta_i\, n_{iM}} - 1 + t_{df_i}^{0.025} \sqrt{\frac{\sum_{j=1}^{n_{iM}} (C_{iMj} - \bar{C}_{iM.})^2}{(n_{iM}-1)\,\theta_i^2\, n_{iM}}}$$

$$\hat{B}_i^{0.975} = \frac{\sum_{j=1}^{n_{iM}} C_{iMj}}{\theta_i\, n_{iM}} - 1 + t_{df_i}^{0.975} \sqrt{\frac{\sum_{j=1}^{n_{iM}} (C_{iMj} - \bar{C}_{iM.})^2}{(n_{iM}-1)\,\theta_i^2\, n_{iM}}}$$

If the bias and precision are assumed constant over k concentrations, the estimates for the bias, the standard error of the bias, and the limits for the 95% confidence interval estimate of the bias over those concentrations are given by:

$$(8)\quad \hat{B} = \frac{\sum_{i=1}^{k}\sum_{j=1}^{n_{iM}} \frac{C_{iMj}}{\theta_i}}{n_{.M}} - 1$$

$$\hat{\sigma}_{\hat{B}} = \sqrt{\frac{\sum_{i=1}^{k}\sum_{j=1}^{n_{iM}} \frac{(C_{iMj} - \bar{C}_{iM.})^2}{\theta_i^2}}{n_{.M}\, df}}$$

$$\hat{B}^{0.025} = \hat{B} + t_{df}^{0.025}\, \hat{\sigma}_{\hat{B}}$$

$$\hat{B}^{0.975} = \hat{B} + t_{df}^{0.975}\, \hat{\sigma}_{\hat{B}}$$

$$\text{where}\quad df = \sum_{i=1}^{k} (n_{iM}-1) \quad \textit{(see text)}$$

$$n_{.M} = \sum_{i=1}^{k} n_{iM}$$

The 95 percent confidence interval limits can be written in terms of the measurements as follows:

$$(8a)\quad \hat{B}^{0.025} = \frac{\sum_{i=1}^{k}\sum_{j=1}^{n_{iM}} \frac{C_{iMj}}{\theta_i}}{n_{.M}} - 1 + t_{df}^{0.025} \sqrt{\sum_{i=1}^{k}\sum_{j=1}^{n_{iM}} \frac{(C_{iMj} - \bar{C}_{iM.})^2}{\theta_i^2\, n_{.M}\, df}}$$

$$\hat{B}^{0.975} = \frac{\sum_{i=1}^{k}\sum_{j=1}^{n_{iM}} \frac{C_{iMj}}{\theta_i}}{n_{.M}} - 1 + t_{df}^{0.975} \sqrt{\sum_{i=1}^{k}\sum_{j=1}^{n_{iM}} \frac{(C_{iMj} - \bar{C}_{iM.})^2}{\theta_i^2\, n_{.M}\, df}}$$

If the studied method precision, S_r, and bias are constant over the concentrations for which the pooled bias estimate is obtained, then the degrees of freedom are as shown; otherwise, the above is less approximate to the extent that S_r and/or the bias varies over the concentrations.

(2) Testing Homogeneity of Bias If the bias varies appreciably from one concentration to another, then the accuracy of the method is almost surely not independent of the concentration and the compliance of the method with the 25% accuracy criterion must be assessed for each concentration. Moreover, unless there is a pattern to the variation of accuracy by concentration which can be statistically inferred, then the information gained from the study of the experimentally selected concentrations may not be useful in predicting method accuracy for other concentrations. However, if it is scientifically reasonable to assume that the accuracy does not vary substantially over the range of concentrations considered in these Guidelines and this assumption passes the following statistical test, then the analysis of accuracy is considerably simplified.

Under the joint hypothesis that the precision and the bias do not vary over the k concentrations, the quantity

$$(9) \quad G = \frac{\sum_{i=1}^{k} n_{iM} (\hat{B}_i - \hat{B})^2/(k-1)}{n_{.M} \hat{\sigma}_{\hat{B}}^2}$$

has a central F distribution with (k-1) and df degrees of freedom. If G is greater than the 95th percentile of such an F-distribution, then the assumption of homogeneity of bias - and thus, of accuracy - over concentrations is contradicted by the data.

b. Concentrations Known from Stoichiometric Determinations for an Experimental Setup (For full details see the second note at the end of this appendix.) Generated concentrations (X/Y g/L) can sometimes be predicted from the amount of material vaporized/aerosolized (X g/min) and the air flow through the generation system (Y L/min). However, small variations in the materials, weighing and measuring errors, mechanical variations, environmental variations (e.g., temperature, humidity, etc.), etc. may produce achieved concentrations which vary from the target or predicted concentrations. The following experimental setup translates these randomlike effects into an increase in the total precision and permits use of the predicted or target concentration as though it were known without error. The "price" is an overestimate of method precision. It is recommended that a statistician be involved in the actual implementation of this design and the analysis of the data.

(1) The Experimental Setup and Terminology This setup is to be used only for the sampling and analysis experiments and not for the recovery experiment. This setup assumes that a concentration can be targeted in advance and realized without bias, i.e., systematic error, by setting up a generation system.

Let the total number of observations planned at the ith concentration be Jn_i, where J is the number of generations at each concentration and n is the number of samplers (measurements) at each generation at the ith concentration. Let Θ_i be the targeted concentration at the ith concentration. The system that, on the basis of stoichiometric theory, will generate a concentration of Θ_i is to be set up (jointly) independently and

run J (for J > 1) times at each concentration.

The Jn_i samplers are to be randomly divided into J equal sets. Each set of samplers is to be uniquely assigned to one of the J independent generations to sample the generated amount which is then analyzed with a correction for recovery.

Let C_{ijl} be the recovery corrected measurement for the lth sampler, l= 1,2,...,n_i, at the jth generation, j= 1,2,...,J, and at the ith concentration, i=1,2,...,k[a]. Note, first, that the n_i samplers are unique to the jth generation at the ith concentration while each generation is, of course, unique to a concentration. Secondly, note that the targeted concentrations, $\{\Theta_i, i=1,2,...,k\}$, are known without error. The statistics and estimators required for method evaluation are given as described in the following pages.

(2) Precision Estimation The estimate of the relative standard deviation for the generation experiment at the ith concentration is given by:

$$(10) \quad \hat{S}_{ri2} = \sqrt{\frac{\sum_{j=1}^{J}\sum_{l=1}^{n_i}(C_{ijl} - \bar{C}_{ij.})^2}{J(n_i - 1)\,\bar{C}_{i..}^2}}$$

$$\text{where } \bar{C}_{ij.} = \sum_{l=1}^{n_i}\frac{C_{ijl}}{n_i} \quad \text{and}$$

$$\bar{C}_{i..} = \sum_{j=1}^{J}\sum_{l=1}^{n_i}\frac{C_{ijl}}{Jn_i}$$

and which has $J(n_i - 1)$ degrees of freedom

for $i = 1,2,\ldots,k$ concentrations.

$\hat{S}_{ri2}$ is used as described in Appendix 2[b].

(3) Bias Estimation The estimated bias for the ith concentration is:

[a] The subscript, M, is omitted here for simplicity; for this setup there is no need for an independent method.

[b] If the variance between generations is so small as to be inconsequential (a term the user must define), then that component can be added to the sum of squares in the numerator under the radical to produce an estimator of S_{ri2} which has larger degrees of freedom. The hypothesis of no generation-dependent component to the total variance can be tested with the data. However, the power of this test may be insufficient to detect a "consequential" between generation component to the variance.

$$(11) \qquad \hat{B}_i = \frac{\bar{C}_{i..}}{\Theta_i} - 1$$

where the $\{\Theta_i\}$ *are the known and targeted concentrations*

and the $\{\bar{C}_{i..}\}$ *are the means of the measurements*

for $i = 1,2,\ldots,k$ *concentrations.*

The estimate of the standard error of the estimated bias at the ith concentration is given by:

$$(12) \qquad \hat{\sigma}_{\hat{B}_i} = \sqrt{\sum_{j=1}^{J} \frac{(\bar{C}_{ij.} - \bar{C}_{i..})^2}{J\,(J-1)\,\Theta_i^2}}$$

which has $(J - 1)$ *degrees of freedom,*

for $i = 1,2,\ldots k$ *concentrations.*

Note: This quantity does not exist if J < 2.

The limits for the 95% confidence interval estimate of the bias at the ith concentration are given by:

$$(13) \qquad \hat{B}_i^{0.025} = \hat{B}_i + t_{df_i}^{0.025}\,\hat{\sigma}_{\hat{B}_i}$$

$$\hat{B}_i^{0.975} = \hat{B}_i + t_{df_i}^{0.975}\,\hat{\sigma}_{\hat{B}_i}$$

where $df_i = J - 1$

The limits can be written in terms of the measurements as:

$$(14) \qquad \hat{B}_i^{0.025} = \frac{\bar{C}_{i..}}{\Theta_i} - 1 + t_{df_i}^{0.025} \sqrt{\frac{\sum_{j=1}^{J} (\bar{C}_{ij.} - \bar{C}_{i..})^2}{J\,(J-1)\,\Theta_i^2}}$$

$$\hat{B}_i^{0.975} = \frac{\bar{C}_{i..}}{\Theta_i} - 1 + t_{df_i}^{0.975} \sqrt{\frac{\sum_{j=1}^{J} (\bar{C}_{ij.} - \bar{C}_{i..})^2}{J\,(J-1)\,\Theta_i^2}}$$

where $df_i = J - 1$

If it is assumed that (see the second note at the end of this appendix for definitions)

$$RSD_{G_i} = \frac{\delta_i}{\Theta_i} = RSD_G, \text{ a constant,}$$

$$(15) \qquad \frac{\sigma_i}{\Theta_i} = (B+1)\, S_{rT}, \text{ a constant, and}$$

$$B_i = B, \text{ a constant, so that}$$

$$S_{ri} = S_{rT}, \text{ a constant,}$$

$$\text{for } i = 1,2,\ldots,k \text{ concentrations,}$$

i.e., that the RSDs of the generation errors, the RSDs of the method errors relative to the target concentrations, the biases of the errors of the study method, and the RSDs of the method errors relative to the study method means are, respectively, homogeneous, over concentrations, then the bias estimate for the entire range of concentrations studied is (note: the homogeneity of the RSDs is required for the confidence interval estimates but not for the point estimates):

$$(16) \quad \hat{B} = \frac{\sum_{i=1}^{k} \omega_i^2 \hat{B}_i}{\sum_{i=1}^{k} \omega_i^2} = \frac{\sum_{i=1}^{k} \frac{\omega_i^2 \bar{C}_{i..}}{\Theta_i}}{\sum_{i=1}^{k} \omega_i^2} - 1$$

$$\text{where } \omega_i^2 = n_i \left[n_i \frac{RSD_G^2}{S_{rT}^2} + 1\right]^{-1}.$$

$$\text{If } n_i = n, \text{ a constant, for } i = 1,2,\ldots k, \text{ then}$$

$$(16a) \quad \hat{B} = \frac{\sum_{i=1}^{k} \hat{B}_i}{k} = \sum_{i=1}^{k} \frac{\bar{C}_{i..}}{k\,\Theta_i} - 1 \ .$$

If the sample size, n_i, is not the same for all concentrations, then see the second note for the determination of the values for the $\{\omega_i\}$.

The estimator for the standard error of the bias estimator is:

$$(17) \quad \hat{\sigma}_{\hat{B}} = \sqrt{\frac{\sum_{i=1}^{k} \omega_i^4 \Theta_i^{-2} \sum_{j=1}^{J} (\bar{C}_{ij.} - \bar{C}_{i..})^2}{\left(\sum_{i=1}^{k} \omega_i^2\right)^2 J\,(J-1)}}$$

$$\text{with } df = k(J-1) \text{ degrees of freedom.}$$

$$\text{If } n_i = n, \text{ a constant, for } i = 1,2,\ldots,k, \text{ then}$$

$$(17a) \quad \hat{\sigma}_{\hat{B}} = \sqrt{\frac{\sum_{i=1}^{k} \Theta_i^{-2} \sum_{j=1}^{J} (\bar{C}_{ij.} - \bar{C}_{i..})^2}{k^2\, J\,(J-1)}}$$

Note: This quantity does not exist if J < 2.

The limits for the 95 percent confidence interval estimate of the bias are:

$$(18) \qquad \hat{B}^{0.025} = \hat{B} + t_{df}^{0.025}\,\hat{\sigma}_{\hat{B}}$$
$$\hat{B}^{0.975} = \hat{B} + t_{df}^{0.975}\,\hat{\sigma}_{\hat{B}}$$

These estimates can be rewritten in terms of the measurements as follows:

$$(19) \quad \hat{B} = \frac{\sum_{i=1}^{k} \frac{\omega_i^2 \bar{C}_{i..}}{\Theta_i}}{\sum_{i=1}^{k} \omega_i^2} - 1 = \frac{\sum_{i=1}^{k} \frac{\omega_i^2 \sum_{j=1}^{J} \bar{C}_{ij.}}{\Theta_i}}{J\sum_{i=1}^{k} \omega_i^2} - 1 = \frac{\sum_{i=1}^{k} \frac{\omega_i^2}{n_i\Theta_i} \sum_{j=1}^{J}\sum_{l=1}^{n_i} C_{ijl}}{J \sum_{i=1}^{k} \omega_i^2} - 1$$

$$\hat{B}^{0.025} = \frac{\sum_{i=1}^{k} \frac{\omega_i^2 \bar{C}_{i..}}{\Theta_i}}{\sum_{i=1}^{k} \omega_i^2} - 1 + t_{df}^{0.025} \sqrt{\frac{\sum_{i=1}^{k} \omega_i^4 \Theta_i^{-2} \sum_{j=1}^{J} (\bar{C}_{ij.} - \bar{C}_{i..})^2}{(\sum_{i=1}^{k} \omega_i^2)^2 \; J\,(J-1)}}$$

$$\hat{B}^{0.975} = \frac{\sum_{i=1}^{k} \frac{\omega_i^2 \bar{C}_{i..}}{\Theta_i}}{\sum_{i=1}^{k} \omega_i^2} - 1 + t_{df}^{0.975} \sqrt{\frac{\sum_{i=1}^{k} \omega_i^4 \Theta_i^{-2} \sum_{j=1}^{J} (\bar{C}_{ij.} - \bar{C}_{i..})^2}{(\sum_{i=1}^{k} \omega_i^2)^2 \; J\,(J-1)}}$$

with $df = k(J-1)$ *degrees of freedom*

(4) Testing Homogeneity of Bias The following statistic can be used to test homogeneity of bias for this case:

$$(20) \quad G = \frac{k\,J\,(J-1) \sum_{i=1}^{k} \omega_i^2\,(\hat{B}_i - \hat{B})^2}{(k-1) \sum_{i=1}^{k} \omega_i^2\,\Theta_i^{-2} \sum_{j=1}^{J} (\bar{C}_{ij.} - \bar{C}_{i..})^2}$$

If $n_i = n$, *a constant, for* $i = 1,2,\ldots,k$, *then*

$$(20a) \quad G = \frac{k\,J\,(J-1) \sum_{i=1}^{k} (\hat{B}_i - \hat{B})^2}{(k-1) \sum_{i=1}^{k} \Theta_i^{-2} \sum_{j=1}^{J} (\bar{C}_{ij.} - \bar{C}_{i..})^2}$$

Under the joint hypothesis that the bias and precision are the same for all concentrations, G has a central F distribution with (k-1) and k(J-1) degrees of freedom. If G is greater than the 95 percentile of such a distribution, then the data contradict the hypothesis of homogeneity of bias over concentrations. Otherwise, that hypothesis is not contradicted. No test is possible if J < 2.

2. *Measured Concentrations Not Known* In the case that the measured concentrations are not known, then an independent method must be used to estimate them. Such a method, ideally, should have no bias and high precision. Replicate samples of the independent method must be collected at each concentration studied. The number of replicates should be at least the same as for the method under evaluation, unless the independent method is less imprecise (e.g., at least 50% smaller RSD) than the studied method. In this case, it would be better to increase the sample size for the studied method and have fewer measurements by the independent method. It would be desirable to have a greater sample size for the independent method if it is more imprecise than the study method. A reasonable goal would be to have sample sizes so that the standard deviation of the bias estimate is as low as possible. For example, as shown in the first note to this appendix, if a single measurement by the study method costs c times that of the independent method and the total cost for samples must less than some budgeted maximum, t, then the ratio of the sample sizes at any concentration, i, should be:

$$(21) \quad \frac{n_{iM}}{n_{iI}} = \frac{S_{ri}}{\sqrt{t}\, S_{rIi}}$$

Estimates of the relative standard deviations will usually have to be used instead of the unknown parameter values.

However, since the studied method observations are needed both to estimate its bias and its S_r, it is better to sacrifice replicates of the independent method than the studied method if the total sample size must be rigidly limited.

There are two sub-cases: a) when the independent and studied methods are not "paired" or blocked together; and b) when "pairing" or blocking is used. The "pairing", which is a form of blocking, need not be 1 to 1 but may be 1 study method to r independent method observations, or - less likely - s study method to r independent observations. There are two important advantages of the "pairing". First, the means of the observations which are paired do not have to have equal variances. The "pairs" or blocks may be formed after the data have been collected to obtain this advantage if all measurements at each concentration were taken under fully randomized conditions. Second, if the blocked measurements are taken under similar conditions (which is planned in advance), the variance of their difference (or the difference of the their logarithms) will be reduced; this yields a more precise estimate of the bias. If the measurements by the two methods are taken under similar conditions, then these measurements will be correlated and treating them as not paired will underestimate the standard error of the bias. The disadvantage is that one degree of freedom for the estimate of the standard error of the bias is lost for each "pair" or block. If the total number of "pairs" for all concentrations is more than 12 and if the S_r of the study method is expected to be constant over concentrations, the loss of degrees of freedom from blocking may not be great.

Pairing 1 study method to r independent method measurements may be used if the precision of the latter is much worse than the former method. The number of replicates of the independent method, r, should be chosen in accordance with Equation 21.

a. Methods Not Paired When the methods are not "paired" or observed in blocks, the bias estimate and the approximate limits of the 95% confidence interval estimate of the bias at the ith concentration are given by (no estimate of the standard error is given because a logarithmic transformation is used as an intermediate step):

$$\hat{B}_i = EXP[\bar{S}_{i.} - \bar{I}_{i.}] - 1$$

$$(22) \quad \hat{B}_i^{0.025} = EXP[\bar{S}_{i.} - \bar{I}_{i.} + t_{df_i}^{0.025} \hat{\sigma}_{\bar{S}_{i.} \pm \bar{I}_{i.}}] - 1$$

$$\hat{B}_i^{0.975} = EXP[\bar{S}_{i.} - \bar{I}_{i.} + t_{df_i}^{0.975} \hat{\sigma}_{\bar{S}_{i.} \pm \bar{I}_{i.}}] - 1$$

$$where \quad \hat{\sigma}_{\bar{S}_{i.} \pm \bar{I}_{i.}} = \sqrt{\frac{\hat{\sigma}_{iS}^2}{n_{iM}} + \frac{\hat{\sigma}_{iI}^2}{n_{iI}}}$$

$$df_i \approx n_{iM}+n_{iI}-2 \quad (see\ text)$$

$$S_{ij} = \ln C_{iMj}\ , \qquad I_{ij} = \ln C_{iIj}$$

$$\bar{S}_{i.} = \frac{\sum_{j=1}^{n_{iM}} S_{ij}}{n_{iM}}\ , \qquad \bar{I}_{i.} = \frac{\sum_{j=1}^{n_{iI}} I_{ij}}{n_{iI}}$$

$$\hat{\sigma}_{iI} = \sqrt{\frac{\sum_{j=1}^{n_{iI}} (I_{ij}-\bar{I}_{i.})^2}{n_{iI} - 1}}$$

$$= The\ Estimated\ Standard\ Deviation\ of \{I_{ij}, j=1, \ldots, n_{iI}\}$$

$$\hat{\sigma}_{iS} = \sqrt{\frac{\sum_{j=1}^{n_{iM}} (S_{ij}-\bar{S}_{i.})^2}{n_{iM} - 1}}$$

$$= The\ Estimated\ Standard\ Deviation\ of \{S_{ij}, j=1, \ldots, n_{iM}\}$$

C_{iMj} = the jth of the n_{iM} measurements on the method under evaluation at the ith concentration;

C_{iIj} = the jth of the n_{I} measurements on the independent method at the ith concentration;

If the variances of the log-transformed measurements of the two methods are known to be equal, then the following "pooled" estimate may be used:

$$\hat{\sigma}_{\bar{S}_{i.} \pm \bar{I}_{i.}} = \sqrt{\frac{\sum_{j=1}^{n_{iM}} (S_{ij} - \bar{S}_{i.})^2 + \sum_{j=1}^{n_{iI}} (I_{ij} - \bar{I}_{i.})^2}{n_{iM} + n_{iI} - 2}}$$

with df_i degrees of freedom.

The notation, "ln", is used for the natural logarithm of its argument while "exp" is used to raise Naperian e to the power of the argument. If the S_r's of the two methods are equal, the relation for the degrees of freedom is nearly exact, otherwise it is approximate (Satterthwaite's approximation might be used instead of the formula shown). This uses the fact that the variance of the logarithms of the observations for a method is approximately equal to the square of the method's S_r.

When the bias and the precision of both the study and independent methods are assumed to be constant over k concentrations and the methods are not "paired" or observed in blocks, the bias estimate and the approximate limits of the 95% confidence interval estimate of the bias pooled over those k concentrations are:

$$\hat{B} = EXP[\bar{S}_{..} - \bar{I}_{..}] - 1$$

$$\hat{B}^{0.025} = EXP[\bar{S}_{..} - \bar{I}_{..} + t_{df}^{0.025} \hat{\sigma}_{\bar{S}_{..} \pm \bar{I}_{..}}] - 1 \qquad (23)$$

$$\hat{B}^{0.975} = EXP[\bar{S}_{..} - \bar{I}_{..} + t_{df}^{0.975} \hat{\sigma}_{\bar{S}_{..} \pm \bar{I}_{..}}] - 1$$

$$\text{where } \hat{\sigma}_{\bar{S}_{..} \pm \bar{I}_{..}} = \sqrt{\frac{\hat{\sigma}_S^2}{n_{.M}} + \frac{\hat{\sigma}_I^2}{n_{.I}}}$$

$$df \approx n_{.M} + n_{.I} - 2k \quad (see\ text)$$

$$S_{ij} = \ln C_{iMj}, \qquad I_{ij} = \ln C_{iIj}$$

$$\bar{S}_{..} = \frac{\sum_{i=1}^{k} \sum_{j=1}^{n_{iM}} S_{ij}}{n_{.M}}, \qquad \bar{I}_{..} = \frac{\sum_{i=1}^{k} \sum_{j=1}^{n_{iI}} I_{ij}}{n_{.I}}$$

$$\hat{\sigma}_S = \sqrt{\frac{\sum_{i=1}^{k} \sum_{j=1}^{n_{iM}} (S_{ij} - \bar{S}_{i.})^2}{n_{.M} - k}}, \qquad \hat{\sigma}_I^2 = \sqrt{\frac{\sum_{i=1}^{k} \sum_{j=1}^{n_{iI}} (I_{ij} - \bar{I}_{i.})^2}{n_{.I} - k}}$$

Testing Homogeneity of Bias The approach is to take the logarithms of the measurements and treat the data as a two-way ANOVA layout in concentrations and methods and test for interaction between concentration and method. This can be accomplished by any standard ANOVA computing software when the numbers of measurements for the two methods are equal or proportional for each concentration. The test statistic is denoted by G and is given by:

Use when $n_{iM} = p\, n_{iI}$ *for every i*
where p is some constant.

$$(24)\quad G = \frac{\sum_{i=1}^{k}\sum_{m=1}^{2} (\bar{V}_{im.} - \bar{V}_{i..} - \bar{V}_{.m.} + \bar{V}_{...})^2 / (k-1)}{\sum_{i=1}^{k}\sum_{m=1}^{2}\sum_{j=1}^{n_{im}} (V_{imj} - \bar{V}_{im.})^2 / (n_{..} - 2k)}$$

$$\text{where } V_{i1j} = \ln(C_{iMj})$$

$$V_{i2j} = \ln(C_{iIj})$$

Under the hypothesis that the bias does not depend on concentration, G has an F distribution with (k-1) and (n.. - 2k) degrees of freedom. If G is less than the 95th percentile for such an F distribution, the hypothesis of homogeneity of bias is not contracted by the data. Otherwise, accuracy should be determined separately for each concentration.

If the numbers of measurements do not satisfy the condition described above, the same approach can be used by using a general linear model approach, such as with SAS PROC GLM[2].

In either of the above cases, the test is approximate if the precision of the study method and the independent method differ substantially. The test can be improved by multiplying the logarithm transformed measurements for each method by the inverse of the RSD of the method. If the precisions are not known, then estimates of the RSDs can be used. This analysis also assumes the homogeneity of the precision of both the study and the independent methods over the k concentrations.

b. Methods Paired This analysis might be used when the user determines that the variances of the study method and the test method are substantially different or when measurements are substantially related to the position of the samplers in the chamber and samplers for the two methods are positioned in proximity by pairs (which are "blocks" technically). The decision might be based on the effects of such differences or chamber position relationships on the final estimates of accuracy or on convenience. When "pairing" or blocking is used, there are n_i "pairs" or blocks at the ith concentration. Each "pair" is treated as a pair of the means for the two methods for the block. Each such "mean" may be only one measurement or be the mean of v measurements where v may vary between the two methods but is constant over blocks within and among concentrations for each method. Thus, each observation in the following is the mean of v observations where v is at least 1. Therefore, C_{iMj} and C_{iIj} are used rather than using $\bar{C}_{iMj.}$ *and* $\bar{C}_{iIj.}$, which involve a fourth subscript, to denote the means over v_M and v_I, respectively, observations.

The bias estimate and the limits of the 95% confidence interval estimate for the bias for the ith concentration are given by:

$$\hat{B}_i = EXP[\bar{S}_{i.} - \bar{I}_{i.}] - 1$$

$$\hat{B}_i^{0.025} = EXP[\bar{S}_{i.} - \bar{I}_{i.} + t_{df_i}^{0.025} \frac{\hat{\sigma}_D(i)}{\sqrt{n_i}}] - 1 \qquad (25)$$

$$\hat{B}_i^{0.975} = EXP[\bar{S}_{i.} - \bar{I}_{i.} + t_{df_i}^{0.975} \frac{\hat{\sigma}_D(i)}{\sqrt{n_i}}] - 1$$

where $$\hat{\sigma}_D(i) = \sqrt{\sum_{j=1}^{n_i} \frac{(D_{ij} - \bar{D}_{i.})^2}{df_i}},$$

$= The\ Estimated\ Standard\ Deviation\ of\ \{D_{ij}\}\ in\ Concentration\ i,$

$$and\ df_i = n_i - 1$$

$$S_{ij} = \ln C_{iMj}\ , \qquad I_{ij} = \ln C_{iIj}\ , \qquad D_{ij} = S_{ij} - I_{ij}$$

$$\bar{S}_{i.} = \frac{\sum_{j=1}^{n_i} S_{ij}}{n_i}\ , \qquad \bar{I}_{i.} = \frac{\sum_{j=1}^{n_i} I_{ij}}{n_i}\ , \qquad \bar{D}_{i.} = \bar{S}_{i.} - \bar{I}_{i.}$$

In the expressions in Equation 25, $\hat{\sigma}_D(i)$ denotes the estimate of the standard deviation of D_{ij} for the ith concentration.

Alternately, these "propagation of errors" approximations to the confidence limits, which are quite satisfactory if $-0.35 < B_i < 0.35$[a], can be used:

$$\hat{B}_i^{0.025} \approx \hat{B}_{i.} + t_{df_i}^{0.025} (\hat{B}_i + 1) \frac{\hat{\sigma}_D(i)}{\sqrt{n_i}} \qquad (25a)$$

$$\hat{B}_i^{0.975} \approx \hat{B}_{i.} + t_{df_i}^{0.975} (\hat{B}_i + 1) \frac{\hat{\sigma}_D(i)}{\sqrt{n_i}}$$

When the methods are paired and if the bias and the precision for the difference, D_{ij}, are assumed to be constant over k concentrations, then the bias estimate and the limits for the 95% confidence interval estimate for the bias for the k concentrations are given by (using the definitions for Equation 25):

[a]Based on the authors' unpublished analysis.

$$\hat{B} = EXP[\bar{S}_{..} - \bar{I}_{..}] - 1$$

$$(26) \quad \hat{B}^{0.025} = EXP[\bar{S}_{..} - \bar{I}_{..} + t_{df}^{0.025} \frac{\hat{\sigma}_D}{\sqrt{n_.}}] - 1$$

$$\hat{B}^{0.975} = EXP[\bar{S}_{..} - \bar{I}_{..} + t_{df}^{0.975} \frac{\hat{\sigma}_D}{\sqrt{n_.}}] - 1$$

where

$df = n_. - k$

$$\hat{\sigma}_D = \sqrt{\frac{\sum_{i=1}^{k}\sum_{j=1}^{n_i}(D_{ij}-\bar{D}_{i.})^2}{df}},$$

The Estimated Standard Deviation of $\{D_{ij}\}$ for all Levels.

Alternately, using a "propagation of errors" approach, these approximations to the confidence limit estimates, which are quite satisfactory if $-0.35 < B < 0.35$[a], might be used:

$$(26a) \quad \hat{B}^{0.025} \approx \hat{B} + t_{df}^{0.025}(\hat{B}+1)\frac{\hat{\sigma}_D}{\sqrt{n_.}}$$
$$\hat{B}^{0.975} \approx \hat{B} + t_{df}^{0.975}(\hat{B}+1)\frac{\hat{\sigma}_D}{\sqrt{n_.}}$$

Testing for Homogeneity of Bias The basic approach is to use the $\{D_{ij}\}$ observations defined for Equation 25 and test for a concentration effect for a one-way layout.

Under the hypothesis that the bias does not vary over the k concentrations and assuming that the precision for the difference, D_{ij}, is constant over the k concentrations, the quantity

$$(27) \quad G = \frac{\sum_{i=1}^{k} n_i(\bar{D}_{i.} - \bar{D}_{..})^2 / (k-1)}{\hat{\sigma}_D^2}$$

has a central F distribution with (k-1) and df degrees of freedom. If G is less than the 95th percentile of such an F distribution, then the assumption of homogeneity of bias is not contradicted by the data. Otherwise, accuracy should be estimated for each concentration.

[a] Based on the authors' unpublished analysis.

3. Using an Independent Estimate of the Bias This is when the estimate of the bias is obtained using, completely or partially, data which is not used to estimate the precision of the method being evaluated. To use the procedures for confidence interval estimation of the accuracy that are described in V of this appendix, it is necessary that the limits of a 95% confidence interval estimate of the bias be obtained as well as a single "point" estimate.

III. TESTING FOR ACCEPTABLE BIAS To test the hypothesis that the method bias is 10% or less, use the 95 percent confidence interval estimate for the bias. If this interval includes any point less than or equal to 0.10 in absolute value, then the bias is acceptable. Otherwise, the hypothesis that the bias is 10% or less is contradicted by the data. This is a two-sided test with a Type I (false positive) error probability of 0.05.

IV. ESTIMATED PRECISION The method S_r is estimated from evaluation experiment data by methods given in Appendix 2 which yield a result designated as $\hat{S}_{rT}$. The limits for a 95% confidence interval estimate of the method precision with a 0.05 pump contribution to the error are given by the following equation adapted from Hald[3]:

$$(28a) \quad Let \quad CV^2_{1-\alpha} \approx \frac{\hat{S}^2_{rT} - (0.05)^2}{\left[1 + U(\alpha)\sqrt{\frac{1}{2df} + \frac{\hat{S}_{rT}{}^2 - (0.05)^2}{n}}\right]^2} + (0.05)^2$$

Then the limits are given as follows:

$$(28b) \quad \begin{aligned} \hat{S}_r^{0.025} &= CV_{1-0.975} \\ \hat{S}_r^{0.975} &= CV_{1-0.025} \end{aligned}$$

Or,

$$(28c) \quad \hat{S}_r^{0.025} = \sqrt{\frac{\hat{S}^2_{rT} - (0.05)^2}{\left[1 + U(0.975)\sqrt{\frac{1}{2df} + \frac{\hat{S}^2_{rT} - (0.05)^2}{n}}\right]^2} + (0.05)^2}$$

$$\hat{S}_r^{0.975} = \sqrt{\frac{\hat{S}^2_{rT} - (0.05)^2}{\left[1 + U(0.025)\sqrt{\frac{1}{2df} + \frac{\hat{S}^2_{rT} - (0.05)^2}{n}}\right]^2} + (0.05)^2}$$

where $\hat{S}_{rT}$ = the estimated method relative standard deviation based on pooling within and possibly between concentrations (see Appendix 2) after a 0.05 pump contribution to the error has been added;

$U(\alpha)$ = the α x 100 percentile of the standard unit normal distribution;
df = the degrees of freedom used to estimate $\hat{S}_{rT}$; and
n = the total sample size for the method under study for the generation samples for all concentrations.

The degrees of freedom, df, for $\hat{S}_{rT}$ are equal, conservatively, to the number of measurements used to obtain the estimate minus the number of concentrations. A less conservative approximation would require the use of Satterthwaite's approximation[4].

If the contribution to the error from the pump is assumed to be any other value than 0.05, then replace all cases that 0.05 is used in Equation 28 with the appropriate value.

V. ESTIMATED ACCURACY The accuracy of a method is the (theoretical) maximum error of a measurement, expressed as the proportion or percentage of the amount being measured without regard for the direction of the error, that is achieved with 0.95 probability by the method. Under the assumption of a normal distribution of measurements by the study and independent methods, accuracy, **A**, is related to the bias, **B**, and the precision, **P**, as follows[6]:

$$(29)\quad 0.95 = \Phi\left(\frac{A - B}{(1 + B) \cdot P}\right) - \Phi\left(\frac{-A - B}{(1 + B) \cdot P}\right),$$

where $\Phi(X) = Pr\{Y \leq X\}$, *and*

Y is a standard normal random variable.

The following describes procedures for the single value and 90 percent (two-sided) confidence interval estimation of accuracy using the estimates of the bias and precision obtained under II and IV. The procedures to obtain the (single value) estimated accuracy use the estimates of the bias, $\hat{B}$ or the $\{\hat{B}_i\}$ (for separate estimates for each concentration), from II and the estimated precision corrected for the pump contribution to error, $\hat{S}_{rT}$, or the $\hat{S}_i$ (for separate estimates by concentration) from Appendix 2.

A. Three Possible Conclusions The estimated accuracy and the 90% confidence interval estimate of the accuracy provide information about the accuracy of the method. These estimates also make possible one of three conclusions about the accuracy of the method relative to the 25% accuracy criterion or any accuracy criterion specifying a different accuracy. The conclusion depends on the relative location of the 90% confidence interval estimate of the accuracy to 25% or whatever accuracy is specified by the AC. The three possibilities and the conclusions associated with them are: (1) if the interval is completely less than 25%, then there is 95% confidence that the method fulfills the AC; (2) if the interval is greater than 25%, then there is 95% confidence that the method does NOT fulfill the AC; and (3) if the interval includes 25%, then there is neither 95% confidence that the method does or does not fulfill the AC[a]. These three conclusions may be given the abbreviated labels of Acceptance, Rejection, and Inconclusive relative to the AC. The third conclusion, the inconclusive case, usually means that more research is required to produce a definite conclusion because the accuracy is close to 25% and/or the sample sizes were too small.

Please note that failure to achieve 95% confidence that the method fulfills an AC does not imply that the method does not fulfill the AC. Nor does failure to achieve 95% confidence that a method does NOT fulfill the AC imply that it does fulfill the AC.

---

[a] This should not be confused with a test of the null hypothesis that the accuracy is 0.25 with a two-sided test. In this case, there are two hypotheses: H1: $I > 0.25$; and H2: $I \leq 0.25$. Note that the point, 0.25, is located in the parameter space specified by the second hypothesis and the two hypotheses completely exhaust the parameter space. One and only one of the two hypotheses can be true. Thus, there is one and only one possible Type I error and its probability is 0.05. Thus, if H1 is rejected for H2, there is 95% confidence that $I \leq 0.25$. Likewise, if H2 is rejected for H1, there is 95% confidence that $I > 0.25$. It is possible that neither hypothesis is rejected, which yields the inconclusive case.

These possibilities are illustrated in Figure 2 for four hypothetical methods labelled A, B, C, and D. Methods A and D are accepted as fulfilling the AC as the 90% confidence interval estimates of their accuracies are less than 25%. That implies that for methods A and D there is at least 95% confidence that the accuracy is less than or equal to 25%. Method B is rejected as not fulfilling the AC as its interval is greater than 25%. Finally, method C illustrates the inconclusive case that the method is neither accepted nor rejected in terms of its fulfillment of the AC. However, the intervals provide more information than that as may be noted in Figure 2.

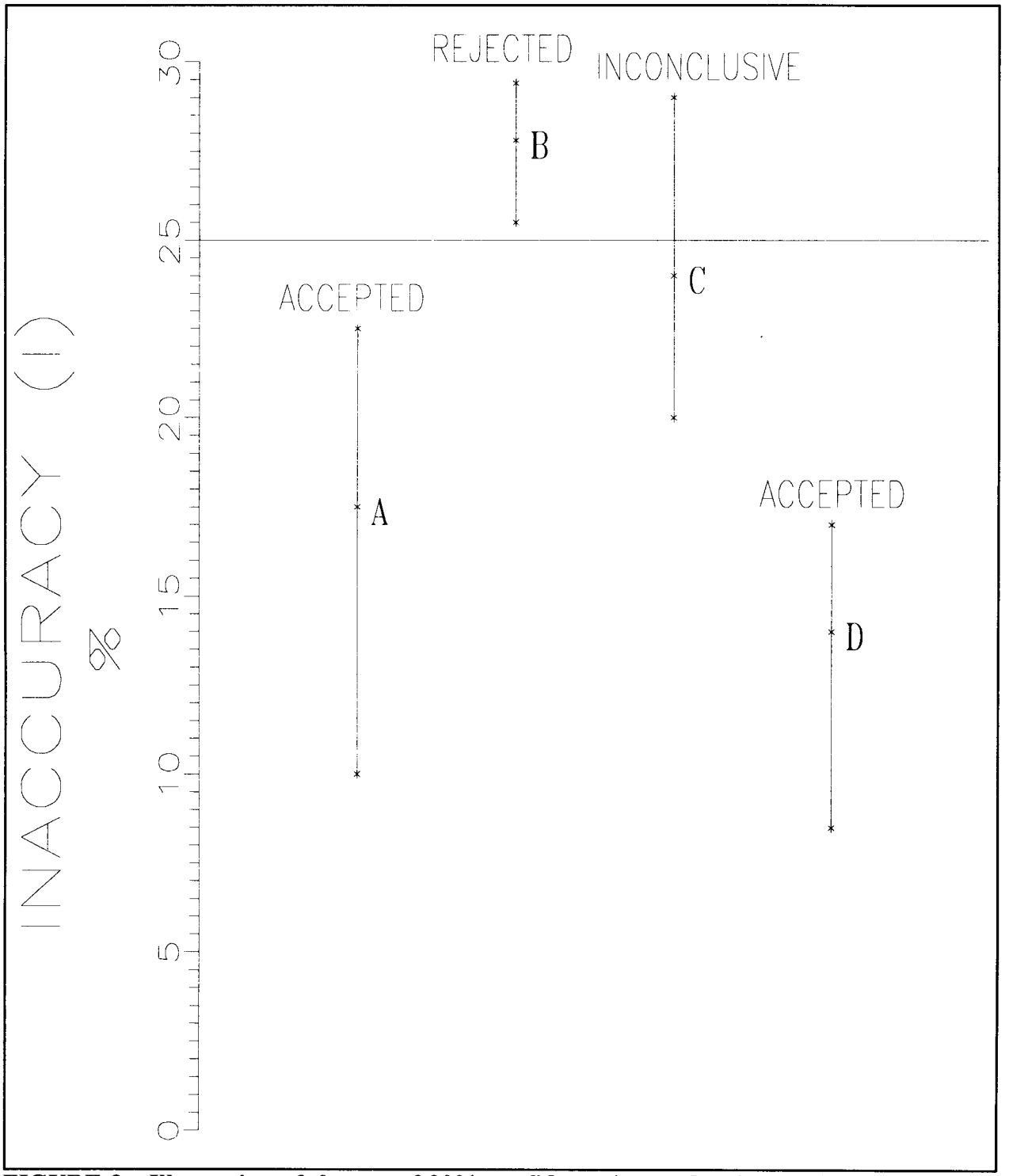


FIGURE 2 - Illustration of the use of 90% confidence interval estimates for three possible conclusions concerning the fulfillment of the 25% Accuracy Criterion using four hypothetical methods labelled A, B, C, D.

The accuracy of method B is at most 30% and may be less than 26%. Method C may be as accurate as method A or as accurate as method B, but it may not be as accurate as method D (although such comparisons must be done with care as the four interval estimates do not hold with 95% joint confidence. Method D is the only method with 95% confidence of an accuracy less than 20%. Either methods A or D may be have an accuracy of only 10%.

B. Two Confidence Interval Construction Procedures_This Appendix presents two alternative procedures for the construction of 90 percent confidence interval estimates of the accuracy which differ considerably in their respective complexity. The first interval estimate procedure is an application of the Bonferroni inequality[5] (the user is advised to note footnote "b" concerning accuracy limit estimates that appears on page one of this appendix). This procedure is described fully in Section V.D.4.

The second procedure, the hyperbolic approximation procedure, for 90 percent confidence interval estimation of accuracy has been developed from hyperbolic approximations to accuracy and accuracy confidence limit estimates obtained from numerical integration[6]. The hyperbolic approximation procedure uses only the estimated (point) estimates of the bias, $\hat{B}$ or the $\{ \hat{B}_i \}$ (for separate estimates for each concentration), from II and the estimated precision corrected for the pump contribution to error, $\hat{S}_{rT}$, or the $\hat{S}_i$ (for separate estimates by concentration) from Appendix 2.

A. Which Procedure to Use for Accuracy Interval Estimation The Bonferroni interval estimation procedure results in conservative limits, that is, the interval estimates that are produced are too wide or the actual confidence for the interval produced is much higher than 90 percent (except in the special cases that either the bias or the precision is known without error). This is because this procedure is based on intervals for the bias and the precision which are to hold simultaneously but the function relating bias and precision to accuracy is not one-to-one. However, this procedure is well-known, general and, within limits, it does not depend on the specific design used to generate the bias and precision estimates. The hyperbolic approximation is simple to use and may be less conservative than the Bonferroni procedure, but the formulas presented in this Appendix are specific to evaluation designs with twelve samplers for each of one, two, three, or four concentrations (a set of formulas for nine samples per concentration is also provided). Formulas for other designs can be developed using the basic procedures used by Song[6]. The recommended choices are as follows in the Table III:

TABLE III - Recommended Accuracy Confidence Interval Construction Procedures

Bias Estimation Case	Section Bias Case Discussed	Recommended Accuracy Interval Estimate Construction Procedure
Bias Estimated: Concentration Assumed Known Without Error	II.B.1.a	Bonferroni (Another hyperbolic approximation can be developed[6]: see a statistician.)
Bias Estimated: Concentration Known Using Special Generation Setup	II.B.1.b	Bonferroni (Another hyperbolic approximation is possible[6]: see a statistician.)
Bias Estimated: Concentration Estimated by Independent Method	II.B.2.a and II.B.2.b	Hyperbolic Approximation
Bias Estimated: All Other Cases	II.B.3	Bonferroni unless a hyperbolic approximation is known to be appropriate[6].

B. **Accuracy Estimate Computing Methods** - The notation for the accuracy estimates is as follows:

$$\begin{aligned} &\hat{A} \quad \textit{is the point estimate,} \\ (30)\quad &\hat{A}^{0.05} \textit{ is the 5 percent confidence statistic, and} \\ &\hat{A}^{0.95} \textit{ is the 95 percent confidence statistic.} \end{aligned}$$

1. *Accuracy Nomogram* The accuracy nomogram, Figure 1, graphs the relationship among bias, precision, and accuracy described in Fischbach, Shulman, and Song[7] and shown in Equation 29. It can be used to obtain a theoretical accuracy, a single value estimate of the accuracy, or an interval estimate using the Bonferroni procedure. The ordinate axis of the graph is method precision while the abscissa is method bias. Accuracy is marked by the curved lines on the graph. By plotting the estimates of any two of these attributes, the estimate for the third one can be obtained. It can be used to obtain a point estimate, $\hat{A}$, for method accuracy by plotting the estimated method bias, $\hat{B}$, as the abscissa and the estimated precision, $\hat{S}_{rT}$, as the ordinate. If, instead, the plotted values are theoretical or target, the output will be the theoretical or target method accuracy. Finally, if the plotted values are a $(1-\alpha_1)$ x 100% confidence statistic for the method bias and a $(1-\alpha_2)$ x 100% confidence statistic for the method S_r, where $\alpha_1 + \alpha_2 = 0.05$, then the output will be a 95% confidence statistic for the accuracy, $A^{0.95}$. The arguments for each statistic using the Bonferroni procedure will be described below.

   The nomogram is used as follows: Plot the point having the bias value as the abscissa and the precision as the ordinate. Determine the position of the plotted point relative to the curved lines which are marked in accuracy percentage units. Note values of accuracy for the two curved lines that the point falls between. Do a rough interpolation to produce the accuracy value.

2. *Computer Algorithm* The computer algorithm is written for PC-SAS[8] and requires the input of the estimated method bias, its standard error, the estimated method precision, and other information described in the listing of the algorithm at the end of this appendix. The result is the estimated accuracy satisfying the relationship to the bias and precision estimates that is described in Equation 29. The output also includes the hyperbolic approximation point estimate and 90 percent confidence interval estimates produced by both procedures for accuracy interval estimation.

   The DOS program, ABCV.EXE[a], will also produce estimates which are exact solutions to Equation 29 using the arguments just described.

3. *Hyperbolic Approximation Formulas* There are separate hyperbolic approximation derived formulas for obtaining an accuracy point estimate and interval limit estimates (or confidence statistics) for one, two, three, and four concentrations.

   a. Accuracy Point Estimation Hyperbolic Approximation Formulas These hyperbolic approximation derived formulas can be used to obtain the estimated accuracy, if the estimates of the bias and the precision are used as arguments, or the theoretical accuracy, if theoretical or target values of the bias and precision are the arguments. They produce an estimated accuracy or a theoretical accuracy which is within 1.1

[a] ABCV.EXE is available from the Division of Physical Sciences and Engineering of NIOSH.

percent of the value obtained by the computer algorithm. They are not recommended to obtain confidence statistics for the accuracy (formulas for that purpose are given below in Section V.D.3.b). The formulas (with adjustment for a 5% pump error) are:

$$\hat{A} = 1.57\,(\hat{B}+1)\sqrt{\hat{S}_{rT}^2 + (0.05)^2} + \sqrt{[0.39\,(\hat{B}+1)]^2\,[\hat{S}_{rT}^2 + (0.05)^2] + \hat{B}^2}\ ,\ \textit{for estimation, and}$$

(31)

$$A \approx 1.57\,(B+1)\sqrt{S_{rT}^2 + (0.05)^2} + \sqrt{[0.39\,(B+1)]^2\,[S_{rT}^2 + (0.05)^2] + B^2}\ ,\ \textit{for theoretical values.}$$

In the expression for $\hat{A}$ in Equation 31, $\hat{S}_{rT}$ includes all sources of error including the pump.

b. Accuracy 90 Percent Confidence Interval Estimation Hyperbolic Approximation Formulas. Formulas have been developed from the hyperbolic approximation to estimate the 5 percent and the 95 percent confidence statistics for accuracy for one, two, three, and four concentrations with respective sample sizes of 9 and 12 each. These are: (Note: these expressions use estimates of $B^* = B + 1$ and not B, etc., and include a 5% pump error adjustment.)

(1) Sample size of 9 per concentration:

(32.9.a) *One concentration*

$$\hat{A}_{1,9}^{0.05} = 1.26\hat{B}^*\sqrt{(\hat{S}_{rT}/1.96)^2+0.05^2} + \sqrt{(0.70\hat{B}^*)^2[(\hat{S}_{rT}/1.96)^2+0.05^2]+(\hat{B}^*-1)^2}$$

$$\hat{A}_{1,9}^{0.95} = 1.80\hat{B}^*\sqrt{(1.83\hat{S}_{rT})^2+0.05^2} + \sqrt{(0.16\hat{B}^*)^2[(1.83S_{rT})^2+0.05^2]+(\hat{B}^*-1)^2}$$

(32.9.b) *Two concentrations*

$$\hat{A}_{2,9}^{0.05} = 1.26\hat{B}^*\sqrt{(\hat{S}_{rT}/1.50)^2+0.05^2} + \sqrt{(0.70\hat{B}^*)^2[(\hat{S}_{rT}/1.50)^2+0.05^2]+(\hat{B}^*-1)^2}$$

$$\hat{A}_{2,9}^{0.95} = 1.80\hat{B}^*\sqrt{(1.49\hat{S}_{rT})^2+0.05^2} + \sqrt{(0.16\hat{B}^*)^2[(1.49S_{rT})^2+0.05^2]+(\hat{B}^*-1)^2}$$

(32.9.c) *Three concentrations*

$$\hat{A}_{3,9}^{0.05} = 1.26\hat{B}^*\sqrt{(\hat{S}_{rT}/1.37)^2+0.05^2} + \sqrt{(0.70\hat{B}^*)^2[(\hat{S}_{rT}/1.37)^2+0.05^2]+(\hat{B}^*-1)^2}$$

$$\hat{A}_{3,9}^{0.95} = 1.80\hat{B}^*\sqrt{(1.37\hat{S}_{rT})^2+0.05^2} + \sqrt{(0.16\hat{B}^*)^2[(1.37S_{rT})^2+0.05^2]+(\hat{B}^*-1)^2}$$

(32.9.d) *Four concentrations*

$$\hat{A}_{4,9}^{0.05} = 1.26\hat{B}^*\sqrt{(\hat{S}_{rT}/1.30)^2+0.05^2} + \sqrt{(0.70\hat{B}^*)^2[(\hat{S}_{rT}/1.30)^2+0.05^2]+(\hat{B}^*-1)^2}$$

$$\hat{A}_{4,9}^{0.95} = 1.80\hat{B}^*\sqrt{(1.31\hat{S}_{rT})^2+0.05^2} + \sqrt{(0.16\hat{B}^*)^2[(1.31S_{rT})^2+0.05^2]+(\hat{B}^*-1)^2}$$

(2) Sample size of 12 per concentration:

(32.12.a) *One concentration*

$$\hat{A}^{0.05}_{1,12} = 1.26\,\hat{B}^*\sqrt{(\hat{S}_{rT}/1.75)^2 + (0.05)^2} + \sqrt{(0.70\,\hat{B}^*)^2[(\hat{S}_{rT}/1.75)^2 + (0.05)^2] + (\hat{B}^*-1)^2}$$

$$\hat{A}^{0.95}_{1,12} = 1.80\,\hat{B}^*\sqrt{(1.65\,\hat{S}_{rT})^2 + (0.05)^2} + \sqrt{(0.16\,\hat{B}^*)^2[(1.65\,\hat{S}_{rT})^2 + (0.05)^2] + (\hat{B}^*-1)^2}$$

(32.12.b) *Two concentrations*

$$\hat{A}^{0.05}_{2,12} = 1.26\,\hat{B}^*\sqrt{(\hat{S}_{rT}/1.40)^2 + (0.05)^2} + \sqrt{(0.70\,\hat{B}^*)^2[(\hat{S}_{rT}/1.40)^2 + (0.05)^2] + (\hat{B}^*-1)^2}$$

$$\hat{A}^{0.95}_{2,12} = 1.80\,\hat{B}^*\sqrt{(1.40\,\hat{S}_{rT})^2 + (0.05)^2} + \sqrt{(0.16\,\hat{B}^*)^2[(1.40\,\hat{S}_{rT})^2 + (0.05)^2] + (\hat{B}^*-1)^2}$$

(32.12.c) *Three concentrations*

$$\hat{A}^{0.05}_{3,12} = 1.26\,\hat{B}^*\sqrt{(\hat{S}_{rT}/1.30)^2 + (0.05)^2} + \sqrt{(0.70\,\hat{B}^*)^2\,[(\hat{S}_{rT}/1.30)^2 + (0.05)^2] + (\hat{B}^*-1)^2}$$

$$\hat{A}^{0.95}_{3,12} = 1.80\,\hat{B}^*\sqrt{(1.31\,\hat{S}_{rT})^2 + (0.05)^2} + \sqrt{(0.16\,\hat{B}^*)^2[(1.31\,\hat{S}_{rT})^2 + (0.05)^2] + (\hat{B}^*-1)^2}$$

(32.12.d) *Four concentrations*

$$\hat{A}^{0.05}_{4,12} = 1.26\,\hat{B}^*\sqrt{(\hat{S}_{rT}/1.25)^2 + (0.05)^2} + \sqrt{(0.70\,\hat{B}^*)^2[(\hat{S}_{rT}/1.25)^2 + (0.05)^2] + (\hat{B}^*-1)^2}$$

$$\hat{A}^{0.95}_{4,12} = 1.80\,\hat{B}^*\sqrt{(1.26\,\hat{S}_{rT})^2 + (0.05)^2} + \sqrt{(0.16\,\hat{B}^*)^2[(1.26\,\hat{S}_{rT})^2 + (0.05)^2] + (\hat{B}^*-1)^2}$$

Note: $\hat{A}^{\alpha}_{k,n}$ *is the* α *x 100% confidence statistic for samples with a sample size of n at each of k concentrations.*

$\hat{B}^* = \hat{B} + 1$, *where* $\hat{B}$ *is the estimate of B obtained by using procedures in Section II;*

$\hat{S}_{rT}$ *is the estimate of* S_{rT} *and;*

The pump error = 0.05.

4. *The Bonferroni Procedure for 90% Confidence Interval Estimates of Method Accuracy* The interval estimate is bounded by $(\alpha_1+\alpha_2)$ x 100% and $(1-\alpha_1-\alpha_2)$ x 100% confidence statistics where $\alpha_1 + \alpha_2 = 0.05$. The $(\alpha_1+\alpha_2)$ x 100% confidence statistic for the accuracy is found as the solution of Equation 29 for **A** using an α_1 x 100% confidence statistic for the bias for **B** and an α_2 x 100% confidence statistic for the precision for **P**. The $(1-\alpha_1-\alpha_2)$ x 100% confidence statistic for the accuracy is found as the solution of Equation 29 for **A** using a $(1-\alpha_1)$ x 100% confidence statistic for the bias for **B** and a $(1-\alpha_2)$ x 100% confidence statistic for the precision for **P**. While any values of α_1 and α_2 which satisfy that requirement and are

independent of the data might be used, it is convenient to choose $\alpha_1 = \alpha_2 = 0.025$ and use the limits of the 95 percent confidence interval estimates for the bias obtained under II, $\hat{B}^{0.025}$ *and* $\hat{B}^{0.975}$, and for the precision obtained under IV, $\hat{S}_r^{0.025}$ *and* $\hat{S}_r^{0.975}$

The two limits of the interval estimate for accuracy, when $\alpha_1 = \alpha_2 = 0.025$, are computed using Computer Algorithm II and are further described as follows:

a. 5% Confidence Statistic for Accuracy, $\hat{A}^{0.05}$ The 5% confidence statistic for accuracy, $\hat{A}^{0.05}$, is obtained by using the 2.5% confidence statistic for the precision, $\hat{S}^{0.025}$, and the following value for bias to solve Equation 29. If the range of the 95% confidence interval estimate for bias, ($\hat{B}^{0.025}$, $\hat{B}^{0.975}$), includes the value of 0, then 0 is the value for bias to be used. Otherwise, the value to use is the 97.5% confidence statistic, $\hat{B}^{0.975}$, if $\hat{B}$ is negative, while it is the 2.5% confidence statistic, $\hat{B}^{0.025}$, if $\hat{B}$ is positive.

b. 95% Confidence Statistic for Accuracy, $\hat{A}^{0.95}$ The 95% confidence statistic for accuracy, $\hat{A}^{0.95}$, is obtained by using the 97.5% confidence statistic for the precision, $\hat{S}^{0.975}$, and, if $\hat{B}$ is negative, the 2.5% confidence statistic for the bias, $\hat{B}^{0.025}$, or, if $\hat{B}$ is positive, the 97.5% confidence statistic for the bias, $\hat{B}^{0.975}$ to solve Equation 29.

REFERENCES

1. Norman L. Johnson and Samuel Kotz, Distributions in Statistics: Continuous Univariate Distributions-1, Houghton Mifflin Co., Boston, 1970, p.54.

2. For example, SAS Institute Inc. SAS/STAT™ User's Guide, Release 6.03 Edition, Cary, NC: SAS Institute Inc., 1988

3. A. Hald, Statistical Theory with Engineering Applications. John Wiley & Sons, New York, 1952. p. 302.

4. For example, F. A. Graybill, An Introduction to Linear Statistical Models: Volume I. McGraw-Hill Book Company, Inc., New York, 1961. pp. 368-370.

5. R. G. Miller, Jr., Simultaneous Statistical Inference, McGraw-Hill Book Company, New York, 1966.

6. Song, R., Bartley, D., Fischbach, T., and Shulman, S., A Numerical Procedure for Finding Confidence Limits, Dept. Health and Human Services, Public Health Service, Centers for Disease Control and Prevention, National Institute for Occupational Safety and Health, Cincinnati, Ohio 45226. [In preparation, 1994.] The 95% confidence statistics for the accuracy were found to have actual confidence coefficients greater than 95% for a range of accuracies from 5% to 25% and for a wide range of biases and precisions consistent with those accuracies using computer simulated data. These results appeared in the presentation by Thomas Fischbach and Ruiguang Song, "Study of the Probability of Analytical Method Acceptance Using a 25% Accuracy Criterion by Computer Simulation," at the 1994 Joint Statistics Meetings, San Francisco, August, 1994. There has also been unpublished analytical verification that the actual confidence coefficients for the 5% and 95% confidence statistics are, respectively, less than 5% and more than 95%. The procedure is conservative but less so than the Bonferroni procedure.

7. T. Fischbach, S. Shulman, and R. Song, Some Statistical Procedures for Sampling and Analytical Method Accuracy Tests and Estimation. Submitted to the American Industrial Hygiene Journal for publication. 1993.

8. See, for example, SAS Institute Inc. SAS® Language Guide for Personal Computers, Release 6.03 Edition. Cary, NC: SAS Institute Inc., 1988.

Note on the Allocation of Sample Sizes between the Study Method and the Independent Method

When the concentration is not known and a sample from an independent method is used to determine the bias of the study method, an important question is how to allocate the sample sizes between the two methods (the study method and the independent method) to obtain an optimal estimate for the method (the study method) bias.

Since for each method the sample sizes are set to be equal for all concentrations, it is sufficient to consider this problem on one concentration only. Let $C_{iM1},...,C_{iMn}$ be a random sample from the study method and $C_{iI1},...,C_{iIm}$ be a random sample from the independent method.

For convenience, the independent method is assumed unbiased. Suppose that measurements from both methods are independently normally distributed with small relative standard deviations. In this case, the measurements are approximately log-normal. Let

$$S_{ij}=\ln(C_{iMj}) \quad for\ j=1,...,n_{iM} \quad and \quad I_{ij}=\ln(C_{iIj}) \quad for\ j=1,...,n_{iI}.$$

Then, $\{S_{ij}\}$ and $\{I_{ij}\}$ are approximately normal with variances given by:

$$Var(S_{ij})=\ln(1+S_{ri}^2)\approx S_{ri}^2$$
$$Var(I_{ij})=\ln(1+S_{rIi}^2)\approx S_{rIi}^2.$$

where S_{ri} and S_{rIi} are the relative standard deviations of the study method and the independent method, respectively.

To estimate the method bias, the following estimator is adopted

$$\hat{B}=e^{\bar{S}-\bar{I}}-1$$

$$where \quad \bar{S}=\frac{1}{n_{iM}}\Sigma\ln C_{iMj} \quad and \quad \bar{I}=\frac{1}{n_{iI}}\Sigma\ln C_{iIj}.$$

The variance of the bias estimator is directly proportional to the variance of $\bar{S}-\bar{I}$ which is given by

$$Var(\bar{S}-\bar{I})=Var(\bar{S})+Var(\bar{I})\approx\frac{S_{ri}^2}{n_{iM}}+\frac{S_{rIi}^2}{n_{iI}}.$$

If the cost of a single measurement by the study method is t times the cost of single measurement by the independent method and the total cost (including the cost of measurement for both methods) is fixed at some amount proportional to a constant c, then sample sizes n_{iM} and n_{iI} will be limited by the condition $tn_{iM}+n_{iI}=c$. Now if S_{ri} and S_{rIi} are known, the optimal allocation of sample sizes n_{iM} and n_{iI} will be determined by minimizing the variance of $\bar{S}-\bar{I}$ under the restriction $tn_{iM}+n_{iI}=c$.

$$\text{Let} \qquad f(n_{iM}) = \frac{a^2}{n_{iM}} + \frac{b^2}{c - t n_{iM}}.$$

$$\text{Then} \qquad f'(n_{iM}) = -\frac{a^2}{n_{iM}^2} + \frac{t b^2}{(c - t n_{iM})^2}.$$

$$\text{Set } f'(n_{iM}) = 0, \text{ and solve for } n_{iM}, \text{ it gives}$$

$$n_{iM} = \frac{ac}{ta + \sqrt{t}b}, \text{ and hence } n_{iI} = \frac{\sqrt{t}bc}{ta + \sqrt{t}b}.$$

$$\text{Therefore,} \qquad \frac{n_{iM}}{n_{iI}} = \frac{a}{\sqrt{t}b}.$$

Replace a by S_{ri} and b by S_{rIi}, then the optimal sample size ratio is given by:

$$\frac{n_{iM}}{n_{iI}} = \frac{S_{ri}}{\sqrt{t}\, S_{rIi}}.$$

which depends on only the relative cost of measurements and the RSD's of the two methods and not on the total cost.

That is, more samples should be made from the method which has lower precision (larger RSD) and lower cost for a single measurement compared to the other method. If the two methods have equal cost for a single measurement, the sample sizes should be proportional to their RSDs.

Note on An Experimental Setup for Using the Stoichiometric Determination of the Concentration as Known without Error

The Experimental Setup This setup is to be used only for the sampling and analysis experiments and not for the recovery experiment. This setup assumes that a concentration can be targeted in advance and realized without bias (i.e., systematic error) by setting up a generation system. The actual concentration realized in repeated independent attempts is assumed to be that targeted subject to random error but that the bias is zero. The assumption of zero bias is practically not testable.

Let the total number of observations planned at the ith concentration be Jn_i, where J is the number of generations at each concentration and n_i is the number of samplers (measurements) at each generation at the ith concentration. Let Θ_i be the targeted concentration at the ith concentration. The system that, on the basis of stoichiometric theory, will generate a concentration of Θ_i is to be independently set up and run J times at each concentration. All preparation of materials and equipment for the generation system must be set up from start to finish for each run as though the other runs had not occurred. All preparation errors must be fully and independently replicated on each run within and among all concentrations.

The Jn_i samplers are to be randomly divided into J equal sets. Each set of samplers is to be uniquely assigned to one of the J independent generations to sample the generated amount which is then analyzed with a correction for recovery.

Let C_{ijl} be the recovery corrected measurement for the lth sampler, l= 1,2,...,n_i , at the jth generation, j= 1,2,...,J, and at the ith concentration, i=1,2,...,k. Note, first, that the n_i samplers are unique to the jth generation at the ith concentration while each generation is, of course, unique to a concentration. Secondly, note that the targeted concentrations, $\{\Theta_i, i=1,2,...,k\}$, are known without error. The statistics and estimators required for method evaluation are given in Section II.B.1.b of this appendix. What follows is a statistical justification for those results.

Statistical Justification To show the statistical justification for the above results it is necessary to precisely state the assumed statistical model that relates actual but unknown generated concentrations to those targeted and relates the method measurements to the generated amounts and the bias:

Let T_{ij} be the actual but unknown generated concentration for the jth generation at the ith concentration for which it is assumed that

$$(N2.1) \quad T_{ij} = \Theta_i + \varepsilon_{ij}$$

$$\text{where } E(\varepsilon_{ij}) = 0,$$

$$VAR(\varepsilon_{ij}) = \delta_i^2, \text{ and}$$

$$\varepsilon_{ij} \sim N(0, \delta_i^2) \text{ (independently)}$$

$$\text{for } j = 1,2,\ldots,J \text{ generation runs within}$$

$$i = 1,2,\ldots,k \text{ concentration levels.}$$

The assumption that $E(\varepsilon_{ij}) = 0$ is true if and only if there is no bias or systematic error in realizing the generation of a concentration of Θ_i by the generation system. This assumption is practically not testable.

The measurements, $\{C_{ijl}\}$, are related to the generated concentrations, $\{T_{ij}\}$, by the following model[1]:

$$(N2.2) \quad C_{ijl} = B_i^* T_{ij} + e_{ijl}$$

$$\text{where } E(e_{ijl}) = 0$$

$$Var(e_{ijl}) = \sigma_i^2$$

$$e_{ijl} \sim N(0, \sigma_i^2) \quad (independently)$$

$$for\ l = 1,2,\ldots,n_i\ samplers\ (measurements)\ within$$

$$j = 1,2,\ldots,J\ generation\ runs\ within$$

$$i = 1,2,\ldots,k\ concentration\ levels.$$

Here, the assumption that $E(e_{ijl}) = 0$ means that corrections to correct all correctable bias have been applied. It is also assumed that the two error terms, i.e., the $\{\varepsilon_{ij}\}$ -- the generation errors -- and the $\{e_{ijl}\}$ -- the measurement errors, are statistically independent.

The models for T_{ij} and C_{ijl} imply that:

$$(N2.3) \quad C_{ijl} = B_i^* \Theta_i + B_i^* \varepsilon_{ij} + e_{ijl}$$

Because the $\{\Theta_i\}$ are known without error, it is clear, by well-known statistics results, that the above expression implies that the expressions in Equation 11 for the estimates of the bias for each specific concentration are the linear estimators that are unbiased with least variance, i.e., that they are best in that sense.

The standard error of the estimator of the bias at the ith concentration is given by the square root of the following:

---

[1] The lth sampler (measurement) is nested within the jth generation run which is nested within the ith concentration level. However, the cumbersome subscript notation for nesting, e.g. $C_{ij_{(i)}l_{(ji)}}$, is not used.

Also, this and the model shown in Equation N2.1 should be adapted for more complex error structures if the latter are more realistic for the method studied.

$$(N2.4)\quad VAR(\hat{B}_i) = \frac{VAR(\bar{C}_{i..})}{\theta_i^2}$$

$$Now,\ VAR(\bar{C}_{i..}) = VAR\left(\sum_{j=1}^{J}\sum_{l=1}^{n_i} C_{ijl}\right) / (J\,n_i)^2$$

$$= \left[\sum_{j=1}^{J}\sum_{l=1}^{n_i}\{VAR(C_{ijl}) = B_i^{*2}\,\delta_i^2 + \sigma_i^2 \quad and \quad \sum_{(j'\neq j\ \vee\ l\neq l')} COVAR(C_{ijl}\ C_{ij'l'})\}\right] (Jn_i)^{-2}$$

$$Now,\ VAR(C_{ijl}) = B_i^{*2}\,\delta_i^2 + \sigma_i^2 \quad and$$

$$COVAR(C_{ijl}\ C_{ij'l'}) = \begin{cases} B_i^2\delta_i^2 & if\ j = j' \\ 0 & if\ j \neq j' \end{cases} \quad so\ that$$

$$VAR(\bar{C}_{i..}) = [\{J\,n_i(B_i^{*2}\,\delta_i^2 + \sigma_i^2)\} + \{J\,n_i(n_i - 1)\,(B_i^{*2}\delta_i^2)\}]\,(Jn_i)^{-2}$$

$$= [\sigma_i^2 + n_i B_i^{*2}\delta_i^2]\ n_i^{-1}\ J^{-1} \quad so\ that$$

$$VAR(\hat{B}_i) = [\sigma_i^2 + n_i B_i^{*2}\delta_i^2]\ n_i^{-1}\ J^{-1}\ \theta_i^{-2}$$

$$= \kappa_i^2\ n_i^{-1}\ J^{-1}\ \theta_i^{-2}$$

$$where\ \kappa_i^2 = \sigma_i^2 + n_i B_i^{*2}\delta_i^2$$

The last expression shows that generation errors, reflected by the $\{\ \delta_i^2\ \}$ multiplied by the number of samplers (measurements) for each generation, can have a marked effect on errors of the bias estimate, although this is moderated by the factors, $\{\ B_i^{*\,2}\ \}$. That expression also indicates that for any total sample size of $Jn_i = q$, the variance of the estimator of $\hat{B}_i$ is at a minimum for $J = q/2$ and $n_i = 2$, i.e., the variance is decreased by increasing the number of generations, J, and decreasing the number of measurements for each generation. On the other hand, the degrees of freedom for the estimator of S_{ri} are at a minimum when $J = q/2$.

To justify the estimates of the standard errors of the bias estimators and the use of the Student t intervals, consider the following set of sums of squares, their distributions, and expected values:

$$(N2.5)\quad SS_{1i} = n_i \sum_{j=1}^{J} (\bar{C}_{ij.} - \bar{C}_{i..})^2$$

$$which\ has\ (J - 1)\ degrees\ of\ freedom = df_i$$

$$E(SS_{1i})\,(J-1)^{-1} = \sigma_i^2 + n_i B_i^{*2}\delta_i^2$$

$$= \kappa_i^2$$

$$= VAR(\hat{B}_i)\ x\ \{known\ constants\}$$

$$for\ i = 1,2,\ldots k\ concentration\ levels.$$

The $\{SS_{1i}\}$ are distributed as mutually independent chi square variables weighted by their expected values. This justifies the estimators given in Equations 12, 13, and 14 for the standard errors and the use of the Student t

intervals for the interval estimates for the bias estimate for each concentration.

The justification for the estimator for the bias for all concentrations begins by assuming that the precision of the generation errors, denoted by RSD_G, and the precision of the study method relative to the targeted concentration are both constant over concentrations:

$$RSD_{G_i} = \frac{\delta_i}{\theta_i} = RSD_G, \text{ a constant, and}$$

$$\frac{\sigma_i}{\theta_i} = B^* S_{rT} = (B + 1) S_{rT}, \text{ a constant} \quad (N2.6)$$

$$\text{for } i = 1,2,\ldots,k \text{ concentrations.}$$

Under these assumptions the model for the measurements can be transformed as follows:

$$\bar{C}_{ij.} = B_i^* \theta_i + B_i^* \bar{e}_{ij} + \bar{e}_{ij.} \text{ by averaging over } n_i \text{ measurements}$$

$$\rightarrow \bar{C}^*_{ij.} = B_i^* + B_i^* e^*_{ij} + e^*_{ij.} \text{ by dividing both sides by } \theta_i$$

where

$$\bar{C}^*_{ij.} = \bar{C}_{ij.}/\theta_i$$

$$e^*_{ij} = e_{ij}/\theta_i$$

$$e^*_{ij.} = \bar{e}_{ij.}/\theta_i$$

$$\rightarrow \bar{C}^*_{ij.} = B_i^* + \varphi^*_{ij}$$

where

$$\varphi^*_{ij} = B_i^* e^*_{ij} + \bar{e}^*_{ij.}.$$

$$\text{Now, } VAR(\varphi^*_{ij}) = B_i^{*2} RSD_G^2 + B^{*2} S_{rT}^2 n_i^{-1}$$

$$= B^{*2} S_{rT}^2 n_i^{-1} \left[n_i \frac{B_i^{*2}}{B^{*2}} \frac{RSD_G^2}{S_{rT}^2} + 1\right]$$

$$= B^{*2} S_{rT}^2 \omega_i^{-2}(B_i^*)$$

$$\text{NOTE: } \omega_i^2(B_i) = n_i\left[n_i \frac{B_i^{*2}}{B^{*2}} \frac{RSD_G^2}{S_{rT}^2} + 1\right]^{-1} \text{ (recall } B_i^* = B_i + 1\text{), while,}$$

$$\text{from Equation 16, } \omega_i^2 = n_i\left[n_i \frac{RSD_G^2}{S_{rT}^2} + 1\right]^{-1}$$

$$\rightarrow (N2.7) \quad \bar{C}^{**}_{ij.} = \omega_i(B_i^*) B_i^* + \varphi^{**}_{ij}, \text{ where}$$

$$\bar{C}^{**}_{ij.} = \omega_i(B_i^*) \bar{C}^*_{ij.},$$

$$\varphi^{**}_{ij} = \omega_i(B_i^*) \varphi^*_{ij}, \text{ and}$$

$$VAR(\varphi^{**}_{ij}) = B^{*2} S_{rT}^2 \text{ for } i = 1,2,\ldots,k \text{ concentration levels.}$$

Equation N2.7 is the model for the regression of the transformed measurements on the $\{ \omega_i(B_i^*) \}$ with the concentration specific biases as their coefficients, where the error term has the identical variance for all the k concentrations. Additionally, Equation N2.7 is the model when the bias varies among the concentrations, i.e., the "full" model. The least squares estimators for the biases under the full model are identical to those given by Equation 11 but the standard errors of these estimators reflect the assumptions stated in Expressions N2.6. The estimators based on the full model are used in what follows here but not for Equation 12.

The justification for the estimator for the bias when homogeneity of bias is assumed, i.e., under

$$(N2.8) \quad B_i = B \quad for\ i = 1,2,\ldots,k\ concentrations$$

$$\rightarrow \quad \omega_i(B^*) = \omega_i \ for\ every\ i.$$

is that the regression model given by Equation N2.7 simplifies to the following "reduced" model:

$$(N2.9) \quad \bar{C}_{ij.}^{***} = \omega_i B^* + \varphi_{ij}^{***} ,$$

$$where \begin{cases} \bar{C}_{ij.}^{***} = \omega_i \bar{C}_{ij.}^{*}, \ and \\ \varphi_{ij}^{***} = \omega_i \varphi_{ij}^{*} . \end{cases}$$

It is straight forward to obtain the estimators given in Equations 16, 17, 18, and 19 from this model for the regression of the transformed measurements on the $\{\omega_i\}$.

The "regression sum of squares," (also termed the "sum of squares due to regression"[2]) and the "error sum of squares," (also termed the "sum of squares about regression"[e]) corresponding to the regression model given by Equation N2.7 are needed to justify the test of homogeneity of bias given by Equations 20 and 20a. While the estimators for the concentration specific biases given in Equation 11 do not depend on the $\{\omega_i(B_i)\}$, both the "regression" and "error" sums of squares under the full model do, and these sums of squares will be defined under the assumption that the bias is homogeneous across concentrations, i.e., that given by Equation N2.8.

hen, the regression and error sums of squares, RSS_F and ESS_F,respectively, are functions of the $\{\omega\}$, as defined in Equation 16, and are given by:

ALIGNL(N2.10)

$$ESS_F = \sum_{i=1}^{k} \omega_i^2 \theta_i^{-2} \sum_{j=1}^{J} \bar{C}ij.^2 - J\sum_{i=1}^{k} \omega_i^2 \hat{B}_i^2$$

$$= \sum_{i=1}^{k} \omega_i^2 \theta_i^{-2} \sum_{j=1}^{J} (\bar{C}_{ij.} - \bar{C}_{i..})^2$$

$$which\ has\ k(J-1)\ degrees\ of\ freedom.$$

$$(N2.11) \quad RSS_F = J\sum_{i=1}^{k} \omega_i^2 \hat{B}_i^2$$

$$which\ has\ k\ degrees\ of\ freedom.$$

The "regression sum of squares," RSS_R, for the reduced model given by Equation N2.9 is required for the justification for the test for homogeneity of bias:

---

[2]See N.R. Draper and H. Smith, Applied Regression Analysis, Second Edition, New York, John Wiley & Sons, Inc., 1966, p. 19.

$$RSS_R = J\,\hat{B}^2 \sum_{i=1}^{k} \omega_i^2$$

$$(N2.12) \qquad = J\,\frac{\left(\sum_{i=1}^{k} \omega_i^2\,\hat{B}_i^2\right)^2}{\sum_{i=1}^{k} \omega_i^2}\,,$$

which has 1 *degree of freedom.*

Finally, from standard analysis of variance procedures, the statistic, G, for test for homogeneity of bias in Equation 20 is the ratio that has RSS_F - RSS_R divided by the resultant degrees of freedom for the numerator and ESS_F divided by its degrees of freedom as the denominator.

Estimation of the $\{\omega_i\}$ Equations 16 to 20 for the estimation of the bias common to all concentrations and the test of the hypothesis of homogeneity of bias over concentrations are defined in terms of the constants $\{\omega_i\}$ that are defined in Equation 16 in terms of the unknown parameters B, RSD_G, and S_{rT} as well as the known sample sizes for the concentrations, the $\{n_i\}$. However, all the variation among the $\{\omega\}$ results from the variation among the $\{n_i\}$ as the unknown parameters are constants. Also, in all cases what is needed is not the values for the $\{\omega_i\}$, but the ratios of the square of each to the sum of the squares of the total array of $\{\omega\}$, i.e.,

$$(N2.13) \qquad \rho_i = \frac{\omega_i^2}{\sum_{i=1}^{k} \omega_i^2}\,.$$

It is clear that when the $\{n_i\}$ are all equal then the $\{\omega\}$ will be and ρ will equal 1/k for i=1,2,...,k. This justifies Equations 16a, 17a, and 20a. Four alternatives are presented below for obtaining values of the $\{\omega_i\}$ when the sample sizes vary by concentration. However, the user should understand that getting the exact values of these constants or the ratios defined above is only important for obtaining the optimal estimator for the bias, the best linear unbiased estimator. If approximate values for these constants are used, the result will usually be an estimator which, while not optimal in the sense given, will be quite satisfactory for most method evaluation studies. The four alternatives are as follows:

1. If the variations among the $\{n_i\}$ are small, treat the $\{\rho\}$ as each equal to 1/k. For example, if four concentrations are used with sample sizes which vary by only 1 replication, e.g., 9, 10, 9, and 9, this would be appropriate.
2. If the variation among the sample sizes is judged by the user to be large, then define the $\{\rho_i\}$ proportional to the $\{n_i\}$.
3. (This alternative is less preferable than the first two.) Use the results in this note to estimate the $\{\rho_i\}$ but subtract at least one degree of freedom from the degrees of freedom for the estimate of the standard error of the bias estimate and both the numerator and denominator of the statistic defined in Equation 20 for testing the homogeneity of bias.
4. (This is not recommended except for research on this question.) Use alternative 3 and an iterative procedure to optimize the estimate of the bias by the choices of the $\{\rho_i\}$.

Other Estimates The $\{\delta_i^2\}$, the set of variances of the concentrations over generations can be estimated. First, a set of sums of squares is needed as follows:

$$(N2.14) \quad SS_{2i} = \sum_{j=1}^{J} \sum_{l=1}^{n_i} (C_{ijl} - \bar{C}_{ij.})^2$$

which has $J(n_i - 1)$ *degrees of freedom.*

$$E(SS_{2i}) / (J(n_i - 1)) = \sigma_i^2$$

Also,

$$\delta_i^2 = (\kappa_i^2 - \sigma_i^2) / (n_i B_i^2)$$

Now, a set of estimators for the { δ_i^2 } can be obtained from the two sets of sums of squares previously defined:

$$(N2.15) \quad \hat{\delta}_i^2 = \left| \frac{SS_{1i}}{(J - 1)} - \frac{SS_{2i}}{J(n_i - 1)} \right| / (n_i \hat{B}_i^2)$$

It should be noted that the three component statistics for the estimators of the $\{\delta_i^2\}$ are statistically independent.

This set of estimators can be used to obtain a set of estimators for the relative standard deviations for the generations as follows:

$$(N2.15) \quad \hat{RSD}_{G_i} = \frac{\hat{\delta}_i}{\theta_i}$$

$$(N2.16) \quad \hat{RSD}_G \approx \sqrt{k^{-1} \sum_{i=1}^{k} \left(\frac{\hat{\delta}_i}{\theta_i} \right)^2}$$

The last expression for the estimate of the relative standard deviation of the generation error assumes that there is homogeneity of that RSD over the k concentrations studied.

Remark: The estimation of the { δ_i^2 } and RSD_G^2 is of practical interest for developing improved generation designs for method evaluation. If the generation errors are usually so small as to be negligible, then the $\{\theta_i\}$ might reasonably be treated as known and there would be no need for J to be greater than 1. On the other hand, if these errors are typically large, it would be advisable to use large values of J. Moreover, research on what factors affect generation errors might become useful for controlling generation errors for method evaluation research and for other uses as well.

This material was taken from K.A. Busch and D.G. Taylor. "Statistical Protocol for the NIOSH Validations Tests," in "Chemical Hazards in the Workplace," G. Choudhary, Ed., American Chemical Society, Washington, D.C., pp. 514-517 (much the same material by K. A. Busch also appears in the appendix to Gunderson and Anderson, et al., the fourth reference). This appendix provides the formulae for performing statistical analyses on the methods evaluation data. The information has been generalized to allow for additional concentration levels and higher numbers of samplers. Where the term Coefficient of Variation (CV) was used, relative standard deviation (S_r) is now used for consistency with the NIOSH Manual of Analytical Methods terminology.

NOTE: The equations in this appendix were originally based on 3 concentrations of six samples each. In the evaluation scheme described in this document, up to 4 concentrations and 12 samplers/concentration are used,[a] and the equations have been modified accordingly. To apply the formulae in this Appendix, the number of concentrations are expressed by the variable "k" and the number of samples by the variable "n".

COMPUTATIONAL FORMULAE FOR STATISTICAL ANALYSIS

This appendix gives the formulae and definitions used in the experimental plan to statistically analyze laboratory data from method evaluation tests. Definitions and symbols are listed below:

Mean = *arithmetic mean or average* ($\bar{x}$), *defined as the sum of the observations divided by the number of observations* (n).

Standard Deviation = *The positive square root of the variance, which in turn is defined as the sum of squares of the deviations of the observations from the mean* ($\bar{x}$) *divided by one less than the number of observations* ($n-1$), *the degrees of freedom.*

$$\text{Standard Deviation} = \sqrt{\frac{\sum_{i=1}^{n} (x_i - \bar{x})^2}{n-1}}$$

S_r = the relative standard deviation (RSD), which is defined as the mean divided by the standard deviation, or

$$S_r = \frac{\text{Std Dev}}{\text{Mean}}$$

Total method precision, S_{rT}, reflects contributions from three components: analytical error (including the effect of using a desorption efficiency factor or adjustment for recovery), sampling error, and pump errors.

The analytical error is determined by analyzing six samplers that have been fortified with pure analyte. The sampling error is based on the analysis of 12 (or n) "generated" samplers that were generated by sampling a

[a] See the discussion on pages 13-14 and the footnotes on page 14.

homogeneous atmosphere containing the analyte at a given concentration level. Sampling error includes pump error if the experiment used personal sampling pumps instead of high precision flow control devices (e.g., critical orifices).

The basic "building blocks" for the estimation of method precision are the estimates from the analytical and generation experiments at each concentration and the pooled estimates from each.

The estimate of RSD for the analytical (fortified) samples is S_{r1i}, and S_{r2i} for the generated samples:

S_{r1i} = S_r (estimated value) for the six analytical samples at the ith concentration (0.1, 0.5, 1, and 2X exposure limit) for the recommended sample volume.

S_{r2i} = S_r (estimated value) for the twelve (or n) generated samples at the ith concentration (0.1, 0.5, 1, and 2X exposure limit) for the recommended sample volume.

$\overline{S_{rx}}$ = pooled S_r: the value derived from the S_r (of a given type, e.g. S_{r1} or S_{r2}) obtained from the analysis of n samples at each of k test concentrations.

The formula for the estimate pooled across concentrations:

$$\overline{S_{rx}} = \sqrt{\sum_{i=1}^{k} \frac{f_i \, (S_{rxi})^2}{f}}$$

where:

f_i = degrees of freedom, equal to number of observations minus one (n_i - 1), at the ith concentration;

k = number of concentrations;

S_{rxi} = relative standard deviation (S_{r1i} or S_{r2i}; x = 1 or 2) of the observations at the ith concentration;

$f = \sum_{i=1}^{k} f_i$;

i = index for the k concentrations.

ANALYTICAL ERROR COMPONENT INCLUDING EFFECTS OF ADJUSTMENTS:

$\overline{S_{r1}}$ = pooled relative standard deviation calculated as above based on data for the (k x n) analytical (spiked) samples (k groups of n samples); and

$\overline{S_{rA+DE}}$ = derived correction to $\overline{S_{r1}}$ including precision error due to the use of the desorption efficiency factor, which is an average of n = 6 values.

$$\overline{S_{rA+\overline{DE}}} = \overline{S_{r1}}\sqrt{7/6} = 1.0801\ \overline{S_{r1}},\quad \textit{when } n = 6, \textit{ and}$$

$$= \overline{S_{r1}}\sqrt{\frac{(n+1)}{n}},\quad \textit{in general.}$$

$\overline{S_{r\overline{A+AMR}}}$ = *corrected* $\overline{S_{r1}}$ *analogous to use of a desorption efficiency factor noted above except that this notation is used where the factor is associated with analytical method recovery (AMR) other than for solid sorbents*

Where,

$$\overline{S_{rA+\overline{AMR}}} = 1.0801\ \overline{S_{r1}},\quad \textit{when } n = 6, \textit{ and}$$

$$= \overline{S_{r1}}\sqrt{\frac{(n+1)}{n}},\quad \textit{in general.}$$

SAMPLING ERROR COMPONENT:

$$\textit{If,}\quad \overline{S_{r2}} > \overline{S_{r1}},\quad \textit{then,}$$

$$\overline{S_{rS}} = \sqrt{(\overline{S_{r2}})^2 - (\overline{S_{r1}})^2}$$

which is the estimate of the component of the precision caused purely by sampling error.

ESTIMATION OF THE TOTAL PRECISION: Since S_{rT} is the result of components caused by sampling error, analytical error adjusted for the error induced by the adjustment for either desorption efficiency or recovery, and pump errors[b], its estimator is:

$$\hat{S}_{rT} = \sqrt{(\overline{S_{rS}})^2 + (\overline{S_{rA+\overline{DE}}})^2 + (S_{rP})^2}$$

Where,

S_{rP} = *relative standard deviation due to the pump error, assumed to be equal to* 0.05, *and*

$\hat{S}_{rT}$ = *estimated relative standard deviation of total procedure which consists of the composite variations in sampling and analysis, desorption efficiency, and the pump error*

The above implies that:

$$\hat{S}_{rT} = \sqrt{(\overline{S_{r2}})^2 - (\overline{S_{r1}})^2 + 1.1667\ (\overline{S_{r1}})^2 + (0.05)^2}$$

Then,

$$\hat{S}_{rT} = \sqrt{(\overline{S_{r2}})^2 + 0.1667\ (\overline{S_{r1}})^2 + (0.05)^2}$$

If $\overline{S_{r2}} \leq \overline{S_{r1}}$, *take* $\overline{S_{rS}} = 0$. *Then, replace* $\overline{S_{r1}}$ *by a pooled estimate* $(\overline{S^*_{r1}})$ *based on* $\overline{S_{r1}}$ *and* $\overline{S_{r2}}$,

$$\overline{S^*_{r1}} = \sqrt{\frac{f_1\ \overline{S_{r1}}^2 + f_2\ \overline{S_{r2}}^2}{f_1 + f_2}}$$

where f_1 *and* f_2 *are the respective f-values used in the denominators of* $\overline{S^2_{r1}}$ *and* $\overline{S^2_{r2}}$. *Thus the equation to be used when* $\overline{S_{r2}} \leq \overline{S_{r1}}$ *is:*

$$\hat{S}_{rT} = \sqrt{1.1667(\overline{S^*_{r1}})^2 + (0.05)^2}$$

The Degrees of Freedom for $\hat{S}_{rT}$ can be conservatively set at the sum of all the observations used for the experiment with generated samples (do not use those for the recovery experiment) minus the number of

---

[b]. Note: this is not the "net" effect due to the pump. The net pump effect is estimated by:

$$\hat{S}_{rT} - \sqrt{(\overline{S_{rS}})^2 + (\overline{S_{rA+\overline{DE}}})^2}$$

concentration levels. A less conservative approximation is based on Satterthwaite's approximation[c] as follows:

$$df \approx \frac{(\sum_{I=1}^{L} w_I S_I^2)^2}{\sum_{I=1}^{L} (w_I S_I^2)^2}$$

where df *is the degrees of freedom for* S_c^2 *and*

$$S_c^2 = \sum_{I=1}^{L} w_I S_I^2 \quad \textit{for}$$

a set of constants $w_1, w_2, \ldots w_L$

and a set of relative standard deviations

$$S_1, S_2, \ldots S_L$$

GRUBB'S TEST[1] for rejection of an observation is applied in order to determine if one of the observations should be rejected as being an outlier. The following equation is used for the test:

$$B_1' = \frac{x - \bar{x}}{s} \text{ or } \frac{\bar{x} - x}{s}$$

where:

x = *observation being tested (most distant the mean)*
$\bar{x}$ = *mean of n observations*
s = *standard deviation based on n-1 degrees of freedom*

For any 6 observations, a value can be rejected if $B_1' \geq 1.94$. For any 9 and 12 observations, B_1' must be greater than 2.32 and 2.55, respectively, to reject a value. The B_1' limit is based on a 1% significance level (i.e., a B_1' value calculated from the data can be expected to exceed 1.94 (or 2.32, or 2.55) only 1% of the time if the observation is a legitimate one conforming to the underlying theory). For validation testing, reject no more than two values in a set of results (4 concentrations) and the two may not be in the same concentration.

Bartlett's Test[2] for equality of $S_{\%i}$'s is applied in order to test the justification for "pooling the relative standard deviations[d]" for any set of n x k generated samples (e.g., n = 6, 9, 12, etc. at each of the k concentrations). The following equation for the chi-square (χ^2), with k-1 degrees of freedom, is used:

---

[c] See, for example, F. A. Graybill, An Introduction to Linear Statistical Models: Volume I. McGraw-Hill Book Company, Inc., New York, 1961, pp. 368-370.

[d] The authors' unpublished analyses show that the distribution of the square of the RSD is closely approximated by the distribution of a variance estimate (from a normal sample) multiplied by a constant -- just as the variance estimate, itself, is distributed as a weighted chi-square.

$$\chi^2 = \frac{f \ln(\overline{S_{rx}})^2 - \sum_{i=1}^{k} f_i \ln(S_{rxi})^2}{1 + \frac{1}{3(k-1)}\left[\left(\sum_{i=1}^{k} \frac{1}{f_i}\right) - \frac{1}{f}\right]}$$

where:

$\overline{S_{rx}}$ = *pooled relative standard deviation of k x n generated samples for x=2 or k X n analytical samples for x=1*

S_{rxi} = *relative standard deviation of n samples at the ith level*

f_i = *degrees of freedom associated with* $(S_{rxi})^2$ *and equal to number of observations at the ith level - 1*

$f = \sum_{i=1}^{k} f_i$

In order to "pass" Bartlett's test at the 2.5% significance level (i.e., not find inequality of RSDs), χ^2 must be less than or equal to 7.38 for three concentrations and 9.35 for four concentrations (χ^2 has k-1 degrees of freedom). The corresponding critical values for the 5% significance level are 5.99 and 7.82, respectively, for three and four concentrations (i.e., k = 3 or k = 4). Note: since the assumption of homogeneity of precision over the tested concentrations is critical not only for the issue of the precision but also for the testing and estimation of a constant bias over concentrations, it is recommended that a 5% test, or no smaller than a 2.5% test, of this assumption be used. See Appendix 1 for the relevance of the assumption of a constant RSD to the testing and estimation of a constant bias.

REFERENCES

1. Grubbs FE, Procedures for Detecting Outlying Observations in Samples, Technometrics, 11:1:1-6, February, 1969.

2. Bartlett MS, Some Examples of Statistical Methods of Research in Agriculture and Applied Biology, J. Roy. Stat. Soc. Suppl. 4, 158-159, 1937.

Taken from: NIOSH/DPSE Quality Assurance Manual (December, 1991)

NATIONAL INSTITUTE FOR OCCUPATIONAL SAFETY AND HEALTH
Standard Operating Procedures for Industrial
Hygiene Sampling and Chemical Analysis

Date Issued: January 24, 1984 **SOP 018**
Date Revised: July 18, 1994

Limits of Detection and Quantitation

I. Introduction - This SOP addresses two measures of analytical capability for individual samples. Limit of Detection (LOD) is a decision point used to decide whether to report a significant analyte signal from the sample. Limit of Quantitation (LOQ) is the smallest amount of analyte which can be measured with precision. These parameters relate to the low-concentration end of the analytical working range and do not provide information about accuracy, precision, or sensitivity at higher concentrations.

[Note: This SOP applies to DPSE's routine chemical analyses on individual samples (excluding elemental analyses by ICP-AES and fiber counting by microscopy). Interpretation of results may be different for replicated factorial experiments; see a statistician for assistance in those cases.]

II. Definitions[1]

A. ***Limit of Detection (LOD)*** is the mass of analyte which gives a mean signal $3\sigma_b$ above the mean blank signal, where σ_b is the standard deviation of the blank signal.

B. ***Limit of Quantitation (LOQ)*** is the lowest mass that can be reported with acceptable precision. LOQ is the **larger** of:

(a) the mass corresponding to the mean blank signal + $10\sigma_b$ (i.e., ± 30% uncertainty), or
(b) the mass above which recovery is ≥ 75%.

III. Analytical Implementation - The following procedure uses the variability of low-level analyte responses, in a matrix approximating that of the samples, as an estimate of σ_b.

[Note: The following approach assumes that σ, the standard deviation of analyte signal, is constant over the narrow range in A.1. below, and that the calibration is linear over this range. Otherwise, consult a statistician for alternate approaches.]

A. ***For each sample set***:

1. Prepare N (five or more) **low-level** calibration standards, **spiked on sampling media**, to cover the range from less than the expected LOD to **no greater than** 10 times the expected LOD. These shall be separately-prepared standards, not simply replicate analyses. A reagent blank, spiked on sampling media, may be used as one of the standards, providing that it produces a positive analytical response.

[Note 1: The more standards used, the better the estimate of LOD. The distribution of N standards for most efficient estimation of the slope is: (N-1)/2 standards slightly below the expected LOD, one standard at 5xLOD, and (N-1)/2 standards at 10xLOD.]

SOP 018

[Note 2: If the standards are not spiked on sampling media, the resulting LOD and LOQ must be identified as "instrumental" values unless they are corrected for recovery and the s_y calculated in A.4. below is known to be equal to that for standards spiked on sampling media.]

[Note 3: These low-level standards for estimation of LOD are not to be confused with the additional, higher-level standards needed to complete the construction of a separate calibration curve for quantitation of field samples.]

2. Analyze the N low-level calibration standards under the same conditions as for the field samples.

[Note: If replicate aliquots of the standards are analyzed, average the results for each standard so that N data points are available for use in the LOD calculation.]

3. Graph the responses of the N low-level calibration standards vs. mass of analyte. Obtain the linear regression equation, $\mathbf{Y = mX + b}$, and the predicted responses ($\hat{Y}_i$) at each analyte mass, X.

[Note: If severe nonlinearity exists in this range, use a nonlinear regression technique[2].]

4. Calculate the standard error of the regression:
$s_y = [\Sigma(\hat{Y}_i - Y_i)^2/(N-2)]^{½}$, where N is the number of data points obtained in step A.2 above.

5. Calculate LOD = $3s_y$/slope = $3s_y$/m. Report the LOD as the **highest** of: (a) calculated LOD, (b) lowest calibration standard, or (c) X-intercept (if regression has a negative Y-intercept).

[Note: If the relative standard deviation of the slope is >0.09, use the Song-Fischbach bias reduced estimator.[3]]

6. Calculate LOQ = 3.33 x LOD, or LOQ = mass above which recovery is ≥ 75%, **whichever is greater**.

B. ***Check the calculated LOD*** against other available data (e.g., visual inspection of the calibration graph, strip chart recording of peaks, background present in the field samples, mass spectrometric data, etc.) to make sure that it is a realistic number. If there are doubts as to whether the LOD can be interpreted as in IV.D. below, so state in the report.

IV. Reporting and Interpretation

A. ***Report sample results*** below the LOD as "Not Detected (ND)".

B. ***Report sample results*** between the LOD and LOQ numerically, to two significant figures, and enclosed in parentheses to emphasize the imprecision of the result.

C. ***Give the calculated values*** of LOD (to one significant figure) and LOQ (to two significant figures) in the report.

D. ***Interpretation***

1. False positives. In the absence of interferences, we know, with ≈99% confidence, that an individual sample giving a signal equal to, or greater than, that of the LOD level **does contain** the analyte. That is, the probability of false positives is ≈1%.

SOP 018

2. False negatives. On the other hand, the probability of false negatives at the LOD level is 50% (i.e., half of samples containing this much analyte will fail to give a detectable signal). More generally, the probability of the analyst detecting the analyte when it is **actually present** varies from <1% when the concentration is <<LOD, to 50% at the LOD, to ≈99% at 2xLOD. If it is necessary for the end-user of the data to operate at a lower rate of false negatives than 50%, they can say, for example, that "At a level of 2xLOD or above, 99% of analyte-containing samples have been detected and reported".

   [Note: If pooled results from replicate samples are examined instead of results from samples taken one-at-a-time, a lower LOD results and the above probabilities are changed; see a statistician for interpretation.]

V. REFERENCES

[1] ACS Subcommittee on Environmental Improvement: Principles of Environmental Analysis. Anal.Chem. 55:2210-2218 (1983).

[2] Burkart, J. A.: General procedures for limit of detection calculations in the industrial hygiene chemistry laboratory. Appl.Ind.Hyg. 1:153-155 (1986).

[3] Song, R. and Fischbach, T.: Estimation of the Ratio σ/β in Linear Calibration (in preparation, 1994).

SOP 018

VI. EXAMPLE - Pentamidine isethionate (Seq. 7292)

A. *Analyze Low Standards* [Steps A.3 and A.4, p.2]

pentamidine isethionate, ng/sample	response, peak area
0.153	**214.8**
0.306	**739.3**
0.615	**1575**
1.23	**3790**
2.46	**6921**
4.92	**11526**
9.84	21712
19.7	40033
39.4	82491

Notes: (Low standards in **bold**)

- Expected LOD = 1 ng.
- No response from field blanks
- Responses are for standard solutions, which were shown to give the same s_y as spiked filters.
- A separate recovery study was performed, giving 75% recovery at 50 ng/sample and 16% recovery at 1 ng/sample.

A linear regression using the above six low standards gives the equation:
Y = 280.9 + 2383.4 X (with s_y = 603.8, slope RSD = 0.062, and correlation coefficient = 0.997).

B. *Calculate LOD* [Step A.5, p.2]

1. LOD = $3s_y$/slope = 3 x 603.8/2383.4 = 0.76 ng per sample

2. At this level, the recovery is approximately 16%, so the recovery-corrected LOD is 0.76 ng/sample divided by 0.16:

LOD = 5 ng/sample

No Song-Fischbach correction is needed since the slope RSD is less than 0.09.

C. *Calculate LOQ* [Step A.6, p.2]

The LOQ is the larger of (a) 3.33 x LOD and (b) the smallest mass with 75% recovery; in this case (b) is larger and:

LOQ = 50 ng/sample

Performance Specification	Experimental Design	Interpretation of Results
ANALYTICAL RECOVERY AND STABILITY Recovery of the analyte from the sampler should be complete and precise (Section III.A. and B.)	Fortify sets of 6 samplers with amounts of analyte equivalent to sampling concentrations of 0.1, 0.5, 1.0 and 2.0 times the lower exposure limit for a minimum of 4 h at 0.01 to 0.20 L/min for sorbent based samplers and 1 to 4 L/min for filter based samplers. If the analyte has a ceiling or short term exposure limit, 6 samplers should be fortified with an amount of analyte adjusted for the shorter sampling time required for this type of exposure limit. If the sampler has a backup section, then a like number of separate backup sections should be fortified with amounts of analyte equivalent to 25% of the amount fortified on the front sections of the samplers. Analyze the samples and backup sections. Recap sample solutions and reanalyze with fresh standards after 1 day storage. Additional sets of 6 samples at each of the 4 levels should be prepared and stored for 7 days under ambient conditions.	The recovery of the analyte from the sampler should be $\leq$ 75%. For a sorbent based sampler, the front section of the medium should be greater than 75% for levels equivalent to sampling 0.1, 0.5, 1.0 and 2.0 times the exposure limit. If recovery varies with analyte loading, results should be graphed as recovery vs loading, so that appropriate correction can be made to sample results. Recovery from the backup section of the sampler should be noted so that appropriate recovery corrections can be applied if there is breakthrough during sampling. The reanalysis of the sample solutions on the second day indicates whether immediate analyses after sample preparation is required or not. Results should agree within 5%. Often when processing many samples, it may be necessary to prepare the samples for analysis in a batch mode. In these instances, samples may not be analyzed for 24 h after preparation. If sample solutions are not stable prior to analysis, analysis must be scheduled as quickly after sample preparation as possible. Results from samples stored for 7 days should compare with samples analyzed within 1 day within experimental error. Discrepancies larger than those expected by experimental error indicate stability problems which should be addressed with additional development effort.

Performance Specification	Experimental Design	Interpretation of Results
SAMPLER CAPACITY The capacity of the sampler for the analyte should be defined so that a maximum recommended sampling time and appropriate sampling rate can be specified. (Section III.E.)	Sampling rates typical for the sampling media selected should be used. Typically these may range from 0.01 - 0.20 L/min for sorbent tube samplers to 1-4 L/min for 37-mm filter cassette samplers. Triplicate samplers should be collected at three different flow rates covering the range discussed above for the particular sampler type. Breakthrough of the analyte through the sampler should be monitored for a period up to 900 min (15 h). Flow rates should be based on accurately calibrated sampling pumps or critical orifices. The generated concentration used for capacity determination should be at least 2 times the highest published exposure limit and verified by an independent method. Sampling should be conducted at temperature extremes (<15 °C and >35 °C) to provide the most severe conditions that might be encountered during sampling. To assess the effect of humidity on capacity, capacity should be determined at both low and high humidities. If a particular humidity level is found to reduce sampler capacity, then that level should be used in all further evaluation experiments. A means is required for the determination of the analyte in the effluent from the sampler. This may involve the use of a backup sampler, continuous monitor or other appropriate means which can provide a measure of analyte concentration in the sampler effluent (ca. 1 - 5% of the influent concentration). If more than 5% of the analyte in the total sampler is found on the backup sampling medium, breakthrough has occurred and the capacity of the sampler has been exceeded. If the analyte is a particulate material and collected by filter, the capacity of the filter is defined by the pressure drop across the sampler or by the loading of the filter.	If the collection process is based primarily on adsorption, the breakthrough time should be proportional to the inverse of the flow rate. This relationship can be checked by plotting breakthrough time versus the inverse of the flow rate. If the resulting plot is a straight line, then this relationship should hold for all flow rates in the flow rate range studied. Some nonlinearity in the plot may be noted due to experimental variability and assumptions made to simplify the relationship of breakthrough time and flow rate. Results from these experimental trials should provide a prediction of the capacity of the sampler at various flow rates and sampling times. If the flow rates and sampling times used in the experiment do not provide for sufficient capacity, a lower flow rate range may have to studied and the experiment repeated. With samplers that use reagents for collection of the analyte, the amount of the reagent in the sampler also will be a limiting factor in the capacity of the sampler, based on the stoichiometry of the reaction. Other factors, such as residence time in the sampler and the kinetics of the reaction of the analyte with the reagent, may affect the capacity of this type of sampler. For filter based samplers, the pressure drop across the sampler should be less than 40 inches of water for less than 2 mg of total loading. The time at which the capacity of the sampler is exceeded is the breakthrough time for a given flow rate. To find the maximum recommended sampling time, the breakthrough time is multiplied by 0.667.
ENVIRONMENTAL PARAMETERS The effect of interference and concentration on method performance should be defined. (Section III.F.)	As a minimum, generated atmospheres of 0.1, 1.0 and 2.0 times the exposure limit should be sampled with 12 samplers per level using a fractional factorial experimental design. If there are known interferences for the method, additional sets of samples should be collected at 0.1 and 2.0 times the exposure limit of the analyte with the interference present at 2 times its exposure limit. If there are no known interferences for the method, other environmental parameters may be studied for their effect on method performance.	In an analysis of variance of the experimental results, the effect of concentration should not be a significant factor in method performance. Recoveries at all levels should not be significantly different. Interferences or other environmental factors should not affect method recovery significantly.

Performance Specification	Experimental Design	Interpretation of Results
SAMPLER STABILITY Samples should be stable for at least 7 days after collection. (Section III.H.)	A concentration of 0.5 times the lowest exposure limit should be sampled with 30 samplers for a minimum of 1/2 the recommended sampling time. The samplers should be randomly divided into 1 group of 12, 1 group of 6, and 4 groups of 3, with the group of 12 analyzed as soon after collection as possible. The other group of 6 should be analyzed after 7 days. The remaining groups of 3 should be analyzed after 10, 14, 21, and 30 days.	If the average analysis results of the group of 6 samplers analyzed on day 7 differs from the day 1 results by more than 10%, sample instability is a problem with the method. Either additional precautions may be required for storage or the method may have to be modified to address this problem. If a plot of recovery vs time (30 days) shows that recovery decreased by more than 10%, the samples are stable only for the amount of time where the recovery has not decreased by 10%. If samples need to be stored for longer periods, more restrictive storage conditions are required. If sample instability is still a problem after remedial actions have been attempted, the method does not meet the sample stability criterion and samples will require immediate analysis that will limit the utility of the method.
PRECISION, BIAS, AND ACCURACY Sampler results should be precise and with little or no bias. (Section III.I.)	Results from previous experiments are used. These include sets of 12 samplers collected at 0.1, 0.5, 1.0 and 2.0 times the exposure limit. If the analyte has a ceiling or short term exposure limit (STEL), 12 samples should be collected at the STEL. The concentrations of all levels should be verified by replicates of an independent method.	Before pooling the estimated relative standard deviations, S_{ri}'s , of the 4 sets of samplers, the homogeneity of the population values should be checked using a test, such as Bartlett's test using 5% or 2.5% significance. If the latter are not homogeneous, the sample set collected at 0.1 x exposure limit should be removed and Bartlett's test recalculated. The estimation of the bias is described in Appendix 1. If the true values of the generated concentration are known, then bias can be estimated by subtracting the mean of the method under study from the generated concentration and dividing the result by the generated concentration at each level. See Appendix 1 for more information. If the true values of the generated concentrations are not known, bias is estimated by one of the procedures described in Appendix 1. Homogeneity of bias between levels should be tested using the appropriate test described in Appendix 1. Estimated bias should be less than $\pm$10% to meet the accuracy criterion. The pooled $\mathbf{S}_{rT}$ for the 3 groups of 12 samplers collected under the Environmental Parameters study and the group of 12 samplers collected under the Sampler Stability experiment (analyzed on Day 0) are tested for homogeneity using the procedure described in Appendix 2. This pooled value is then used to estimate precision of the method. Appendix 1 describes procedures to be used under general conditions for testing whether the method does or does not fulfill the accuracy criterion. If the method does not fulfill the accuracy criterion (with 95 percent confidence), then the set of samples collected at 0.1 x exposure limit should be excluded to determine if the method will meet the accuracy criterion for three levels using procedures in Appendix 1. For the STEL measurements, the S_{rT} and bias estimates for the 12 samplers should be obtained and analyzed following procedures in Appendix 1 to determine whether or not the method fulfills the accuracy criterion with 95 percent confidence.

Performance Specification	Experimental Design	Interpretation of Results
FIELD EVALUATION Performance of the method is evaluated in field tests. (Section III.K.)	Both the collection of area samples and personal samples should be included in the field evaluation of the method. Area samples should provide an estimate of field precision and bias. Personal samples may confirm these values and also provide a means to assess the utility of the method. A statistical study design should be prepared based on the variability of the method and the statistical power required to observed differences between the independent method and the method under evaluation. A minimum of 20 pairs of the method under study and an independent method should be used for personal sampling. Placement of the samplers on the workers should be random to prevent the biasing of results due to the "handedness" of the worker. Workers sampled should be in areas where both low and high concentrations of the analyte may be present. Sets of a minimum of 6 area samplers paired with independent methods should be placed in areas of low, intermediate and high analyte concentration. If the atmosphere sampled is not homogeneous, precautions may have to be taken to ensure that all samplers are exposed to the same concentrations.	Field precision and bias of the area sampler results of the method under study should compare with laboratory evaluation results, provided that precautions have been taken to ensure that all samplers have been exposed to the same homogeneous atmosphere. This can be done by using field exposure chambers, such as those described in the literature. Differences in precision and bias can be investigated using either Student's t-test or analysis of variance. Sources of variation should be studied and corrections implemented where necessary. Evaluation of personal sampler results should be done cautiously, since observable differences may be due to work practices or other situations which are beyond the control of the method.

A two-level factorial experimental design can be used to estimate the significant factors and factor interactions which affect method performance. Factorial designs can cover a number of factors, however, as the number of factors increases, the experiment becomes more complicated. The number of experimental trials associated with a factorial experiment are expressed as two raised to the "nth" power or 2^n, where n represents the number of factors. For example, when n=3, the number of trials is 8. A generic format for the factorial experimental design is shown in Table 1. The plus (+) and minus (-) signs indicate that a factor is held either at a high or low value.

Table 2 is used to show how the factor effects and interactions are calculated.[a] The response values for each trial are placed next to the trial number. In experiments where concentration is one of the factors, it is necessary to normalize the data for the two concentrations used. A typical way to do this is to express the data as percent recovery. In each factor and interaction column, the response values with the "+" are added and the total is listed as Sum +. In a similar manner the response values associated with the "-" are added and listed as Sum -. The Sum - values for each column are subtracted from the corresponding Sum + values and the results listed as column Differences. The Difference value for each column is then divided by the number of plus signs in that column to get the estimated Effect values. These Effect values are representative of the factor or interaction effects. The minimum significant factor value (SF_{min}) is calculated using the following equation:

$$SF_{min} = ts\sqrt{2/mk}$$

Where:
t = Student's t value (95% two-tail probability) for the degrees of freedom in s
m = number of "+" in the column
s = pooled standard deviation of a single response observation
k = number of replicates, i.e., samplesize, for each trial.

If the Effect value for a given column is greater than the SF_{min} value, that factor or interaction is statistically significant at the 95% level. This means that this factor or interaction does exert influence on the experimental results. The information on factor and interaction effects gained from these experiments can be used to construct a linear model to predict this influence on results ($Y_{predicted} = k_1 x_1 + ...k_n x_n + k_{12} x_1 x_2 ....$, where k_n refers to the Effect value for a factor or interaction). However, at this point a statistician should be consulted to further interpret results and suggest other statistical designs. For example, it would be advisable to get confidence interval estimates for each coefficient in the model and for the value predicted by the fitted model for various values of the independent variables. A sample calculation is shown in Table 3.

With two replications per trial (i.e., k = 2) a standard deviation of 4.46 with 8 degrees of freedom was calculated. Each sum is the sum of eight values, four from each replicate. The corresponding two-sided t value is 2.31. Based on this value, the following calculation of SF_{min} is:

$$SF_{min} = 2.31 \times 4.46 \sqrt{\frac{2}{4 \times 2}} = 5.15$$

This indicates that factors x_1, x_2, x_3 and interaction of x_2 and x_3 are significant at the 95% confidence level, since the Effect values for these columns exceed the SF_{min} value.

---

[a] This illustrates the process. However, it is probably much more efficient and informative to use a statistics computer software package with an analysis of variance module for these calculations. Most statisticians can provide assistance on this.

TABLE 1 - Two level factorial design patterns for two or three factors. (The bold line defines the 2^2 design.)

Trial #	Factor 1 (x_1)	Factor 2 (x_2)	Factor 3 (x_3)
1	-	-	-
2	+	-	-
3	-	+	-
4 (2^2)	+	+	-
5	-	-	+
6	+	-	+
7	-	+	+
8 (2^3)	+	+	+

TABLE 2 - Calculation format including factor interactions for 2^n (n=2,3) factorial experimental designs. (The bold line defines the 2^2 design.)

Trial #	Results	x_1	x_2	x_1 x_2	x_3	x_1 x_3	x_2 x_3	x_1 x_2 x_3
1	+	-	-	+	-	+	+	-
2	+	+	-	-	-	-	+	+
3	+	-	+	-	-	+	-	+
4	+	+	+	+	-	-	-	-
5	+	-	-	+	+	-	-	+
6	+	+	-	-	+	+	-	-
7	+	-	+	-	+	-	+	-
8	+	+	+	+	+	+	+	+
Sum +								
Sum -								
Difference								
Effect								

TABLE 3 - Sample calculation for a 2^3 factorial experiment.

Trial #	Results	x_1	x_2	$x_1 x_2$	x_3	$x_1 x_3$	$x_2 x_3$	$x_1 x_2 x_3$
1	10	-	-	+	-	+	+	-
2	33	+	-	-	-	-	+	+
3	17	-	+	-	-	+	-	+
4	43	+	+	+	-	-	-	-
5	95	-	-	+	+	-	-	+
6	127	+	-	-	+	+	-	-
7	155	-	+	-	+	-	+	-
8	178	+	+	+	+	+	+	+
Sum +	658	381	393	326	555	332	376	323
Sum -	0	277	265	332	103	326	282	335
Difference	658	104	128	-6	452	6	94	-12
Effect	82.25	26	32	-1.5	113	1.5	23.5	-3

COMPUTER ALGORITHM I - Compute the accuracy of an analytical method from its bias, B, and precision, S_{rT}.

This algorithm, written for PC-SAS[a] provides both an exact and approximate solution for the accuracy as a function of the bias, B ([method mean - quantity measured]/quantity measured), and the precision, S_{rT} (method standard deviation/method mean). If parametric or population values for the bias and precision are entered, the accuracy will be a parametric or population value, A. If the inputs are estimates, the output will be an estimate of the accuracy, A. The exact solution is an iterative solution of Equation 29. The approximate solution is the solution of the hyperbolic approximation given by Equation 31. Use Computer Algorithm II, listed below, not this one, to obtain confidence statistic for the accuracy.

PC-SAS[a] code for Computer Algorithm I follows:

```
options linesize=80;
 title1 'Compute Accuracy with Given Bias and CV';
 title2 'Input Bias and CV or Their Estimates, and Pump Error';
 title3 'Output Accuracy or Its Estimate';
 title4 'OPTIONAL USER TITLE: REPLACE THIS WITH DESIRED TITLE ';
 title5 'OPTIONAL USER SUB-TITLE                              ';
data input;
 Bias= 0.03;                          *** Bias or Bias Estimate ;
 CV  = 0.07;                          *** CV or CV Estimate ;
 pump= 0.05;                          *** The Pump Error       ;

* data input;                         ***                      ;
*  input Bias CV pump;                   ***                   ;
* cards;                              *** For Multiple Inputs ;
* enter data here;                     ***                     ;
* ;                                  ***                      ;

data accuracy; set input;

     CVp=sqrt(CV*CV+pump*pump);              *** CV with Pump Error ;
     TRSD=(Bias+1)*CVp;

    itern=0; low=0.0; high=1.0;                    ***          ;
ITER:itern+1; Accuracy=(low+high)/2.0;                 *** Accuracy  ;
q=probnorm((Bias+Accuracy)/TRSD)-probnorm((Bias-Accuracy)/TRSD);** Iterate  ;
    if q<0.95 then low=Accuracy; else high=Accuracy;      *** Algorithm ;
    if abs(q-0.95)>0.00001 & itern<50 then go to ITER;    ***          ;

                              *** Accuracy (Exact Solution) ;

  HyperA=1.57*TRSD+sqrt((0.39*TRSD)**2+Bias*Bias);

                    *** HyperA = Hyperbolic Approximation of Accuracy ;

 keep Bias CV pump Accuracy HyperA;

       *** Print Input (Bias, CV, pump) and Output (Accuracy and HyperA) ;
```

[a] For example, SAS Institute Inc. SAS® Language Guide for Personal Computers, Release 6.03 Edition. Cary, NC:SAS Institute Inc., 1988.

```
data out; set accuracy;
  file print;
  put // @5 'The Inputs:'
// @10 'Bias or Bias Estimate'          @52 'Bias'    @62 '=' @65 Bias
// @10 ' CV  or  CV  Estimate'          @52 ' CV'     @62 '=' @65 CV
// @10 'The Pump Error'                 @52 'pump'    @62 '=' @65 pump
     // @5 'The Outputs:'
// @10 'The Exact Calculation of Accuracy' @52 'Accuracy' @62 '=' @65 Accuracy
// @10 'The Hyperbolic Approximation of Accuracy' @52 'HyperA'
                                        @62 '=' @65 HyperA;
run;
```

COMPUTER ALGORITHM II - Estimate the Accuracy of an Analytical Method from Estimates of its Bias, B, AND Precision, S_{rT}.

This algorithm, written for PC-SAS[b], provides 95% confidence interval estimates of the bias, B, and the precision, S_{rT}, and a 90% confidence interval estimate of analytical method accuracy, A. This algorithm also displays the results showing whether there is 95% confidence that the method satisfies the accuracy criterion requiring 25% accuracy, or there is 95% confidence that the method does not fulfill the 25% accuracy criterion, or the results are inconclusive. These algorithms use the procedures discussed in Appendix 8.

Computer Algorithm II is presented as three examples with different input data and using different procedures. To use Computer Algorithm II, the user would enter the data appropriate to the procedures that are to be used and replacing the input values for any other data with 0.0.

The presentation of Computer Algorithm II is followed by the "LOG" and "OUTPUT" window outputs for an actual execution of the three examples.

The places where the user must enter or change data input are clearly marked by the comment "/*INSERT YOUR VALUE*/". The type of input required, whether a bias or precision estimate, a degrees of freedom, or a sample size, is specified on the same line as the required input.

"II.B.2" refers to sub-sub-section 3 of sub-section B of section II of Appendix 8 for when an independent method is used to estimate the generated concentration. If this is the case, the required input is the mean of the logarithms of the study method observations, the mean of the logarithms of the independent observations, and a standard deviation which depends on whether Equation 23 is used, enter "stdIS" for the value for $\hat{\sigma}_{\bar{s}_{..} \pm \bar{I}_{..}}$, or Equation 26 is used, enter "stdD" for the value of $\hat{\sigma}_D$. Only one of the latter inputs should be nonzero. "II.B.2 IS not USED" refers to all other cases.

```
*****EXAMPLE 1: II.B.2 IS not USED.  neither EQUATION 23 nor 26 APPLY;
options linesize=80;
  title1 'Compute  5% and 95% Confidence Statistics for Accuracy.';
  title2 'Input Bias estimate, STD and DF of Bias Estimate,';
  title3 'SRT estimate, DF of SRT Estimate, and the Total Sample Size.';
  title4 'Output Confidence Statistics for Bias, SRT, and Accuracy.';
```

[b] For example, SAS Institute Inc. SAS® Language Guide for Personal Computers, Release 6.03 Edition. Cary, NC:SAS Institute Inc., 1988.

```
 title5 'BON05,BON95-Bonferroni Estimates. HYP05,HYP95-Hyperbolic Estimates';
 TITLE8 ' N                                                    N';
 TITLE7 'EXAMPLE 1: II.B.2 IS not USED.  neither EQUATION 23 nor 26 APPLY';

data input;
**** STANDARD BIAS INPUT (ALL BUT II.B.2: MUST BE 0.0 IF II.B.2   ;
 Be  =0.03  /* INSERT YOUR VALUE */ ;          *** Bias Estimate        ;
 stdb=0.04  /* INSERT YOUR VALUE */ ;          *** STD of Bias Estimate  ;
 dfb1 =30   /* INSERT YOUR VALUE */ ;          *** DF of Bias Estimate   ;
**** ALTERNATE BIAS INPUT FOR II.B.2:MUST BE 0.0 IF II.B.2 NOT USED   ;
 SBAR =  0.0  /* INSERT YOUR VALUE */; *** MEAN OF LOGS OF STUDY METHOD OBS.;
 IBAR =  0.0  /* INSERT YOUR VALUE */; *** MEAN OF LOGS OF INDEP METHOD OBS.;
 stdIS=  0.0  /* INSERT YOUR VALUE */; *** SEE EQUATION 23, IF RELEVANT, OR ;
 stdD =  0.0  /* INSERT YOUR VALUE */; *** SEE EQUATION 26, IF RELEVANT.    ;
 dfb2 =  0.0  /* INSERT YOUR VALUE */; *** FROM EQU 23 OR EQU 26, THE DF    ;
**** INPUT THE PRECISION ESTIMATES  ;

 SRTe =0.07 /* INSERT YOUR VALUE */ ;          *** SRT Estimate         ;
 dfSRT=15   /* INSERT YOUR VALUE */ ;          *** DF of SRT Estimate    ;
 n   =18    /* INSERT VALUE */; *** Total Sample Size for the Study Method ;
 pump=0.05;                                  *** The Pump Error       ;

* data input;                                 ***                ;
*  input Be stdb dfb1 SRTe dfSRT n pump;          ***              ;
* cards;                                     *** For Multiple Inputs ;
* enter data here;                            ***                ;
data conlimit; set input;
length  casenm $   16  ;                                  ;

if dfb1   <=  0.0   then do ;
   Be   =   .  ;
   stdb =   .  ;
   dfb1  =   .  ;
   if stdis <= 0.0 then stdis  =  0  ;
   if stdd  <= 0.0 then stdd   =  0  ;
   stdb      =  stdis + stdd/sqrt(n)  ;
   if stdis <= 0.0 then stdis  =  .  ;
   if stdd  <= 0.0 then stdd   =  .  ;
   case =   2  ;
   casenm  = 'II.B.2 USED'   ;
  end;
if dfb2   <=  0.0  then do  ;
   sbar  = .  ;
   ibar  = .  ;
   stdis = .  ;
   stdd  = .  ;
   dfb2  = .  ;
   case  = 1  ;
   CASENM = 'STANDARD ' ;
 end;

 SRT025=SRTe/(1+probit(0.975)*sqrt(1/2/dfSRT+SRTe*SRTe/n)); *** 2.5% Statistic for SRT ;
 SRT025=sqrt(SRT025*SRT025+pump*pump);                  *** Plus Pump Error  ;
```

```
 SRT975=SRTe/(1+probit(0.025)*sqrt(1/2/dfSRT+SRTe*SRTe/n)); *** 97.5% Statistic for SRT;
 SRT975=sqrt(SRT975*SRT975+pump*pump);                    *** Plus Pump Error  ;

if case = 1 then do;
 B025=Be+tinv(0.025,dfb1,0)*stdb;                  *** 2.5% Statistic for Bias;                ;
 B975 =Be+tinv(0.975,dfb1,0)*stdb;                 *** 97.5% Statistic for Bias;
   end;

if case = 2  then  do  ;
  Be   =  exp(sbar - ibar)  - 1.0  ;
  b025 =  exp(sbar - ibar + tinv(0.025,dfb2,0)*stdb)  - 1.0 ;
  b975 =  exp(sbar - ibar + tinv(0.975,dfb2,0)*stdb)  - 1.0 ;
  dfb1  =  dfb2  ;
  end;

 Blow=0; if Be>0 and B025>0 then Blow=B025;
       if Be<0 and B975<0 then Blow=B975;

TRSD1=(Blow+1)*SRT025;

      itern=0; low=0.0; high=1.0;                          ***          ;
ITER1: itern+1; BON05=(low+high)/2.0;                          *** Accuracy  ;
q=probnorm((Blow+BON05)/TRSD1)-probnorm((Blow-BON05)/TRSD1);  *** Iterate   ;
     if q<0.95 then low=BON05; else high=BON05;              *** Algorithm ;
     if abs(q-0.95)>0.00001 & itern<50 then go to ITER1;     ***          ;
                    *** BON05 = Bonferroni Estimate for 5% Statistic ;

  if 11<=dfSRT<22 then c05=1.75+(1.40-1.75)*(dfSRT-11)/11;
  if 22<=dfSRT<33 then c05=1.40+(1.30-1.40)*(dfSRT-22)/11;
  if 33<=dfSRT<44 then c05=1.30+(1.25-1.30)*(dfSRT-33)/11;
  if 44<=dfSRT    then c05=1.25;
  TRSD05=(Be+1)*sqrt((SRTe/c05)**2+pump*pump);
  HYP05=1.26*TRSD05+sqrt((0.70*TRSD05)**2+Be*Be);
                    *** HYP05 = Hyperbolic Estimate for 5% Statistic ;
 Bhigh=abs(Be)+tinv(0.975,dfb1,0)*stdb;
 TRSD2=(Bhigh+1)*SRT975;

      itern=0; low=0.0; high=1.0;                          ***          ;
ITER2: itern+1; BON95=(low+high)/2.0;                          *** Accuracy  ;
q=probnorm((Bhigh+BON95)/TRSD2)-probnorm((Bhigh-BON95)/TRSD2);*** Iterate   ;
     if q<0.95 then low=BON95; else high=BON95;              *** Algorithm ;
     if abs(q-0.95)>0.00001 & itern<50 then go to ITER2;     ***          ;
                    *** BON95 = Bonferroni Estimate for 95% Statistic ;

  if 11<=dfSRT<22 then c95=1.65+(1.40-1.65)*(dfSRT-11)/11;
  if 22<=dfSRT<33 then c95=1.40+(1.31-1.40)*(dfSRT-22)/11;
  if 33<=dfSRT<44 then c95=1.31+(1.26-1.31)*(dfSRT-33)/11;
  if 44<=dfSRT    then c95=1.26;
  TRSD95=(Be+1)*sqrt((c95*SRTe)**2+pump*pump);
  HYP95=1.80*TRSD95+sqrt((0.16*TRSD95)**2+Be*Be);
                    *** HYP95 = Hyperbolic Estimate for 95% Statistic ;
```

```
  length BONpass $ 9;
  BONpass='Uncertain'; if BON05>0.25 then BONpass='NO';
                  if BON95<0.25 then BONpass='YES';
  length HYPpass $ 9;
  HYPpass='Uncertain'; if HYP05>0.25 then HYPpass='NO';
                  if HYP95<0.25 then HYPpass='YES';
   if  case = 2 then stdb  =  .  ;

  keep Be stdb dfb1 SRTe dfSRT n pump sbar ibar stdis stdd dfb2 case
      B025 SRT025 BON05 HYP05 B975 SRT975 BON95 HYP95 BONpass HYPpass
       casenm;

data out; set conlimit;
  file print;
  put /// @5 'The Inputs:'
// @10 'Bias Estimate'                    @55 'Be'    @62 '=' @65 Be
  / @15 'The case is '                   @55 'Case'  @62 '=' @65 casenm
  / @15 'Standard Deviation of Bias Estimate' @55 'stdb'  @62 '=' @65 stdb
  / @15 'Degrees of Freedom of Bias Estimate' @55 'dfb'   @62 '=' @65 dfb1
  / @15 'Mean of logs of Study method obs.  ' @55 'sbar'  @62 '=' @65 sbar
  / @15 'Mean of logs of Indep method obs   ' @55 'ibar'  @62 '=' @65 ibar
  / @15 'Stdis from Equation 23            ' @55 'stdis' @62 '=' @65 stdis
  / @15 'StdD from  Equation 26            ' @55 'stdD'  @62 '=' @65 stdd
// @10 'Precision Estimate'                  @55 'SRTe'  @62 '=' @65 SRTe
  / @15 'Degrees of Freedom of SRT Estimate'  @55 'dfSRT' @62 '=' @65 dfSRT
// @10 'Total Sample Size of the Test Method' @55 'n   '  @62 '=' @65 n
 / @10 'The Pump Error'                  @55 'pump'  @62 '=' @65 pump
     /// @5 'The Outputs:'
//                        @40 '95% Confidence Interval for Bias'
// @10 ' 2.5% Confidence statistic for Bias'      @55 'B025'  @62 '=' @65 B025
 / @10 '97.5% Confidence statistic for Bias'      @55 'B975'  @62 '=' @65 B975
//                        @40 '95% Confidence Interval for Precision'
// @10 ' 2.5% Confidence statistic for SRT'       @55 'SRT025' @62 '=' @65 SRT025
 / @10 '97.5% Confidence statistic for SRT'       @55 'SRT975' @62 '=' @65 SRT975
// @10 'The Bonferroni Approach:'
 /                         @40 '90% Confidence Interval for Accuracy'
 // @15 ' 5% Confidence statistic for Accuracy'   @55 'BON05' @62 '=' @65 BON05
  / @15 '95% Confidence statistic for Accuracy'   @55 'BON95' @62 '=' @65 BON95
   / @17 'Fulfill the Accuracy Criterion ?'  @55 BONpass
// @10 'The Hyperbolic Approach:'
 /                         @40 '90% Confidence Interval for Accuracy'
 // @15 ' 5% Confidence statistic for Accuracy'   @55 'HYP05' @62 '=' @65 HYP05
  / @15 '95% Confidence statistic for Accuracy'   @55 'HYP95' @62 '=' @65 HYP95
   / @17 'Fulfill the Accuracy Criterion ?'  @55 HYPpass;
run;

*****EXAMPLE 2: II.B.2.b IS USED.  EQUATION 26 APPLIES;
*****                                          ;
options linesize=80;
  title1 'Compute  5% and 95% Confidence Statistic for Accuracy.';
  title2 'Input Bias estimate, STD and DF of Bias Estimate,';
  title3 'SRT estimate, DF of SRT Estimate, and the Total Sample Size.';
  title4 'Output Confidence statistic for Bias, SRT, and Accuracy.';
```

```
 title5 'BON05,BON95-Bonferroni Estimates. HYP05,HYP95-Hyperbolic Estimates';
 TITLE8 ' N                                                    N';
 TITLE7 'EXAMPLE 2: II.B.2.b IS USED.  EQUATION 26 APPLIES';

data input;
**** STANDARD BIAS INPUT (ALL BUT II.B.2: MUST BE BLANK IF II.B.2 IS USED;
****                                                          ;
 Be  = 0.0    /* INSERT YOUR VALUE */ ;          *** Bias Estimate        ;
 stdb=  0.0   /* INSERT YOUR VALUE */ ;          *** STD of Bias Estimate  ;
 dfb1 = 0.0   /* INSERT YOUR VALUE */ ;          *** DF of Bias Estimate   ;
****                                                          ;
**** ALTERNATE BIAS INPUT FOR II.B.2:MUST BE BLANK IF II.B.2 NOT USED   ;
****                                                          ;
SBAR = 6.00 /* INSERT YOUR VALUE */;  *** MEAN OF LOGS OF STUDY METHOD OBS.;
IBAR = 6.0296    /* INSERT YOUR VALUE */;  *** MEAN OF LOGS OF INDEP METHOD OBS.;
stdIS=  0.0  /* INSERT YOUR VALUE */;  *** SEE EQUATION 23, IF RELEVANT, OR ;
stdD =  0.2   /* INSERT YOUR VALUE */;  *** SEE EQUATION 26, IF RELEVANT. ;
dfb2 =   18   /* INSERT YOUR VALUE */;  *** FROM EQU 23 OR EQU 26, THE DF ;
****                                                          ;
**** INPUT THE PRECISION ESTIMATES  ;
****                                                          ;

 SRTe =0.07 /* INSERT YOUR VALUE */ ;          *** SRT Estiamte          ;
 dfSRT=15   /* INSERT YOUR VALUE */ ;          *** DF of SRT Estimate     ;
 n   =18    /* INSERT VALUE */; *** Total Sample Size for the Study Method ;
 pump=0.05;                        *** The Pump Error        ;

* data input;                                    ***                  ;
*  input Be stdb dfb1 SRTe dfSRT n pump;                ***                  ;
* cards;                                        *** For Multiple Inputs ;
* enter data here;                               ***                  ;
* ;                                             ***                  ;
data conlimit; set input;
length  casenm $   16  ;                                    ;

if dfb1   <=  0.0   then do ;
   Be   =   .  ;
   stdb =   .  ;
   dfb1  =   .  ;
   if stdis <= 0.0 then stdis  =  0  ;
   if stdd  <= 0.0 then stdd   =  0  ;
   stdb      =  stdis + stdd/sqrt(n)  ;
   if stdis <= 0.0 then stdis  =  .  ;
   if stdd  <= 0.0 then stdd   =  .  ;
   case =   2  ;
   casenm  = 'II.B.2 USED'   ;
  end;
if dfb2   <=  0.0   then do  ;
   sbar  = .  ;
   ibar  = .  ;
   stdis = .  ;
   stdd  = .  ;
   dfb2  = .  ;
```

```
  case  = 1  ;
  CASENM = 'STANDARD ' ;
 end;

 SRT025=SRTe/(1+probit(0.975)*sqrt(1/2/dfSRT+SRTe*SRTe/n)); *** 2.5% Statistic for SRT ;
 SRT025=sqrt(SRT025*SRT025+pump*pump);                      *** Plus Pump Error  ;
 SRT975=SRTe/(1+probit(0.025)*sqrt(1/2/dfSRT+SRTe*SRTe/n)); *** 97.5% Statistic for SRT;
 SRT975=sqrt(SRT975*SRT975+pump*pump);                      *** Plus Pump Error  ;
if case = 1 then do;
 B025=Be+tinv(0.025,dfb1,0)*stdb;                  *** 2.5% Statistic for Bias;                ;
 B975 =Be+tinv(0.975,dfb1,0)*stdb;                 *** 97.5% Statistic for Bias;
   end;
if case = 2  then  do  ;
  Be   =  exp(sbar - ibar)  - 1.0  ;
  b025 =  exp(sbar - ibar + tinv(0.025,dfb2,0)*stdb)  - 1.0 ;
  b975 =  exp(sbar - ibar + tinv(0.975,dfb2,0)*stdb)  - 1.0 ;
  dfb1  =  dfb2  ;
  end;

 Blow=0; if Be>0 and B025>0 then Blow=B025;
       if Be<0 and B975<0 then Blow=B975;

TRSD1=(Blow+1)*SRT025;

      itern=0; low=0.0; high=1.0;                          ***          ;
ITER1: itern+1; BON05=(low+high)/2.0;                      *** Accuracy  ;
q=probnorm((Blow+BON05)/TRSD1)-probnorm((Blow-BON05)/TRSD1);  *** Iterate   ;
     if q<0.95 then low=BON05; else high=BON05;            *** Algorithm ;
     if abs(q-0.95)>0.00001 & itern<50 then go to ITER1;   ***          ;
                    *** BON05 = Bonferroni Estimate for 5% Statistic ;
  if 11<=dfSRT<22 then c05=1.75+(1.40-1.75)*(dfSRT-11)/11;
  if 22<=dfSRT<33 then c05=1.40+(1.30-1.40)*(dfSRT-22)/11;
  if 33<=dfSRT<44 then c05=1.30+(1.25-1.30)*(dfSRT-33)/11;
  if 44<=dfSRT    then c05=1.25;
  TRSD05=(Be+1)*sqrt((SRTe/c05)**2+pump*pump);
  HYP05=1.26*TRSD05+sqrt((0.70*TRSD05)**2+Be*Be);
                    *** HYP05 = Hyperbolic Estimate for 5% Statistic ;
 Bhigh=abs(Be)+tinv(0.975,dfb1,0)*stdb;
 TRSD2=(Bhigh+1)*SRT975;
      itern=0; low=0.0; high=1.0;                          ***          ;
ITER2: itern+1; BON95=(low+high)/2.0;                      *** Accuracy  ;
q=probnorm((Bhigh+BON95)/TRSD2)-probnorm((Bhigh-BON95)/TRSD2);*** Iterate   ;
     if q<0.95 then low=BON95; else high=BON95;            *** Algorithm ;
     if abs(q-0.95)>0.00001 & itern<50 then go to ITER2;   ***          ;
                    *** BON95 = Bonferroni Estimate for 95% Statistic ;
  if 11<=dfSRT<22 then c95=1.65+(1.40-1.65)*(dfSRT-11)/11;
  if 22<=dfSRT<33 then c95=1.40+(1.31-1.40)*(dfSRT-22)/11;
  if 33<=dfSRT<44 then c95=1.31+(1.26-1.31)*(dfSRT-33)/11;
  if 44<=dfSRT    then c95=1.26;
  TRSD95=(Be+1)*sqrt((c95*SRTe)**2+pump*pump);
  HYP95=1.80*TRSD95+sqrt((0.16*TRSD95)**2+Be*Be);
                    *** HYP95 = Hyperbolic Estimate for 95% Statistic ;
  length BONpass $ 9;
```

```
   BONpass='Uncertain'; if BON05>0.25 then BONpass='NO';
                        if BON95<0.25 then BONpass='YES';
   length HYPpass $ 9;
   HYPpass='Uncertain'; if HYP05>0.25 then HYPpass='NO';
                        if HYP95<0.25 then HYPpass='YES';
    if  case = 2 then stdb  =  .  ;
   keep Be stdb dfb1 SRTe dfSRT n pump sbar ibar stdis stdd dfb2 case
        B025 SRT025 BON05 HYP05 B975 SRT975 BON95 HYP95 BONpass HYPpass
         casenm;
data out; set conlimit;
  file print;
  put /// @5 'The Inputs:'
// @10 'Bias Estimate'                    @55 'Be'    @62 '=' @65 Be
  / @15 'The case is '                    @55 'Case'  @62 '=' @65 casenm
  / @15 'Standard Deviation of Bias Estimate' @55 'stdb'  @62 '=' @65 stdb
  / @15 'Degrees of Freedom of Bias Estimate' @55 'dfb'   @62 '=' @65 dfb1
  / @15 'Mean of logs of Study method obs.  ' @55 'sbar'  @62 '=' @65 sbar
  / @15 'Mean of logs of Indep method obs   ' @55 'ibar'  @62 '=' @65 ibar
  / @15 'Stdis from Equation 23             ' @55 'stdis' @62 '=' @65 stdis
  / @15 'StdD from  Equation 26             ' @55 'stdD'  @62 '=' @65 stdd
// @10 'Precision Estimate'                  @55 'SRTe'   @62 '=' @65 SRTe
  / @15 'Degrees of Freedom of SRT Estimate'  @55 'dfSRT'  @62 '=' @65 dfSRT
// @10 'Total Sample Size of the Test Method' @55 'n   '  @62 '=' @65 n
 / @10 'The Pump Error'                     @55 'pump'  @62 '=' @65 pump
     /// @5 'The Outputs:'
//                        @40 '95% Confidence Interval for Bias'
// @10 ' 2.5% Confidence Statistic for Bias'     @55 'B025'  @62 '=' @65 B025
 / @10 '97.5% Confidence Statistic for Bias'     @55 'B975'  @62 '=' @65 B975
//                        @40 '95% Confidence Interval for Precision'
// @10 ' 2.5% Confidence Statistic for SRT'      @55 'SRT025' @62 '=' @65 SRT025
 / @10 '97.5% Confidence Statistic for SRT'      @55 'SRT975' @62 '=' @65 SRT975
// @10 'The Bonferroni Approach:'
 /                      @40 '90% Confidence Interval for Accuracy'
 // @15 ' 5% Confidence Statistic for Accuracy'  @55 'BON05' @62 '=' @65 BON05
  / @15 '95% Confidence Statistic for Accuracy'  @55 'BON95' @62 '=' @65 BON95
   / @17 'Fulfill the Accuracy Criterion ?'  @55 BONpass
// @10 'The Hyperbolic Approach:'
 /                      @40 '90% Confidence Interval for Accuracy'
 // @15 ' 5% Confidence Statistic for Accuracy'  @55 'HYP05' @62 '=' @65 HYP05
  / @15 '95% Confidence Statistic for Accuracy'  @55 'HYP95' @62 '=' @65 HYP95
   / @17 'Fulfill the Accuracy Criterion ?'  @55 HYPpass;
run;

*****EXAMPLE 3: II.B.2.a IS USED.  EQUATION 23 APPLIES;
*****                                          ;
options linesize=80;
  title1 'Compute  5% and 95% Confidence Statistic for Accuracy.';
  title2 'Input Bias estimate, STD and DF of Bias Estimate,';
  title3 'SRT estimate, DF of SRT Estimate, and the Total Sample Size.';
  title4 'Output Confidence Statistics for Bias, SRT, and Accuracy.';
  title5 'BON05,BON95-Bonferroni Estimates. HYP05,HYP95-Hyperbolic Estimates';
  TITLE8 ' N                                              N';
  TITLE7 'EXAMPLE 3: II.B.2.a IS USED.  EQUATION 23 APPLIES';
```

```
data input;
**** STANDARD BIAS INPUT (ALL BUT II.B.2: MUST BE BLANK IF II.B.2 IS USED) ;
****                                                                       ;
 Be  = 0.0    /* INSERT YOUR VALUE */ ;        *** Bias Estimate        ;
 stdb= 0.0    /* INSERT YOUR VALUE */ ;        *** STD of Bias Estimate ;
 dfb1 = 0.0   /* INSERT YOUR VALUE */ ;        *** DF of Bias Estimate  ;
****                                                                       ;
**** ALTERNATE BIAS INPUT FOR II.B.2:MUST BE BLANK IF II.B.2 NOT USED   ;
****                                                                       ;
 SBAR = 7.00 /* INSERT YOUR VALUE */; *** MEAN OF LOGS OF STUDY METHOD OBS.;
 IBAR = 7.22  /* INSERT YOUR VALUE */; *** MEAN OF LOGS OF INDEP METHOD OBS.;
 stdIS=  0.055  /* INSERT YOUR VALUE */; *** SEE EQUATION 23, IF RELEVANT, OR ;
 stdD =  0.0  /* INSERT YOUR VALUE */; *** SEE EQUATION 26, IF RELEVANT. ;
 dfb2 =  18   /* INSERT YOUR VALUE */; *** FROM EQU 23 OR EQU 26, THE DF ;
****                                                                       ;
**** INPUT THE PRECISION ESTIMATES  ;
****                                                                       ;
 SRTe =0.0995 /* INSERT YOUR VALUE */ ;        *** SRT Estiamte         ;
 dfSRT=15    /* INSERT YOUR VALUE */ ;         *** DF of SRT Estimate   ;
 n  =18    /* INSERT VALUE */; *** Total Sample Size for the Study Method ;
 pump=0.05;                    *** The Pump Error       ;

* data input;                                  ***                ;
*  input Be stdb dfb1 SRTe dfSRT n pump;              ***                ;
* cards;                                     *** For Multiple Inputs ;
* enter data here;                            ***                ;
* ;                                          ***                ;
data conlimit; set input;
length  casenm $  16  ;                                  ;

if dfb1  <=  0.0   then do ;
   Be  =  . ;
   stdb =  . ;
   dfb1  =  . ;
   if stdis <= 0.0 then stdis  =  0  ;
   if stdd  <= 0.0 then stdd   =  0  ;
   stdb      =  stdis + stdd/sqrt(n)  ;
   if stdis <= 0.0 then stdis  =  . ;
   if stdd  <= 0.0 then stdd   =  . ;
   case =  2  ;
   casenm  = 'II.B.2 USED'   ;
  end;
if dfb2  <=  0.0  then do  ;
   sbar  = .  ;
   ibar  = .  ;
   stdis = .  ;
   stdd  = .  ;
   dfb2  = .  ;
   case  = 1  ;
   CASENM = 'STANDARD ' ;
 end;
```

```
 SRT025=SRTe/(1+probit(0.975)*sqrt(1/2/dfSRT+SRTe*SRTe/n)); *** 2.5% Statistic for SRT ;
 SRT025=sqrt(SRT025*SRT025+pump*pump);                    *** Plus Pump Error  ;
 SRT975=SRTe/(1+probit(0.025)*sqrt(1/2/dfSRT+SRTe*SRTe/n)); *** 97.5% Statistic for SRT;
 SRT975=sqrt(SRT975*SRT975+pump*pump);                    *** Plus Pump Error  ;

if case = 1 then do;
 B025=Be+tinv(0.025,dfb1,0)*stdb;                  ***  2.5% Statistic for Bias;              ;
 B975 =Be+tinv(0.975,dfb1,0)*stdb;                 *** 97.5% Statistic for Bias;
   end;

if case = 2  then  do  ;
  Be   =  exp(sbar - ibar)  - 1.0  ;
  b025 =  exp(sbar - ibar + tinv(0.025,dfb2,0)*stdb)  - 1.0 ;
  b975 =  exp(sbar - ibar + tinv(0.975,dfb2,0)*stdb)  - 1.0 ;
  dfb1  =  dfb2  ;
  end;

 Blow=0; if Be>0 and B025>0 then Blow=B025;
       if Be<0 and B975<0 then Blow=B975;

TRSD1=(Blow+1)*SRT025;

      itern=0; low=0.0; high=1.0;                        ***          ;
ITER1: itern+1; BON05=(low+high)/2.0;                        *** Accuracy  ;
q=probnorm((Blow+BON05)/TRSD1)-probnorm((Blow-BON05)/TRSD1);  *** Iterate   ;
     if q<0.95 then low=BON05; else high=BON05;              *** Algorithm ;
     if abs(q-0.95)>0.00001 & itern<50 then go to ITER1;     ***          ;
                        *** BON05 = Bonferroni Estimate for 5% Statistic ;

  if 11<=dfSRT<22 then c05=1.75+(1.40-1.75)*(dfSRT-11)/11;
  if 22<=dfSRT<33 then c05=1.40+(1.30-1.40)*(dfSRT-22)/11;
  if 33<=dfSRT<44 then c05=1.30+(1.25-1.30)*(dfSRT-33)/11;
  if 44<=dfSRT    then c05=1.25;
  TRSD05=(Be+1)*sqrt((SRTe/c05)**2+pump*pump);
  HYP05=1.26*TRSD05+sqrt((0.70*TRSD05)**2+Be*Be);
                        *** HYP05 = Hyperbolic Estimate for 5% Statistic ;

 Bhigh=abs(Be)+tinv(0.975,dfb1,0)*stdb;
 TRSD2=(Bhigh+1)*SRT975;

      itern=0; low=0.0; high=1.0;                        ***          ;
ITER2: itern+1; BON95=(low+high)/2.0;                        *** Accuracy  ;
q=probnorm((Bhigh+BON95)/TRSD2)-probnorm((Bhigh-BON95)/TRSD2);*** Iterate   ;
     if q<0.95 then low=BON95; else high=BON95;              *** Algorithm ;
     if abs(q-0.95)>0.00001 & itern<50 then go to ITER2;     ***          ;
                        *** BON95 = Bonferroni Estimate for 95% Statistic ;

  if 11<=dfSRT<22 then c95=1.65+(1.40-1.65)*(dfSRT-11)/11;
  if 22<=dfSRT<33 then c95=1.40+(1.31-1.40)*(dfSRT-22)/11;
  if 33<=dfSRT<44 then c95=1.31+(1.26-1.31)*(dfSRT-33)/11;
  if 44<=dfSRT    then c95=1.26;
  TRSD95=(Be+1)*sqrt((c95*SRTe)**2+pump*pump);
  HYP95=1.80*TRSD95+sqrt((0.16*TRSD95)**2+Be*Be);
```

```
                    *** HYP95 = Hyperbolic Estimate for 95% Statistic ;

  length BONpass $ 9;
  BONpass='Uncertain'; if BON05>0.25 then BONpass='NO';
                  if BON95<0.25 then BONpass='YES';
  length HYPpass $ 9;
  HYPpass='Uncertain'; if HYP05>0.25 then HYPpass='NO';
                  if HYP95<0.25 then HYPpass='YES';

   if  case = 2 then stdb  =  .  ;
  keep Be stdb dfb1 SRTe dfSRT n pump sbar ibar stdis stdd dfb2 case
      B025 SRT025 BON05 HYP05 B975 SRT975 BON95 HYP95 BONpass HYPpass
       casenm;

data out; set conlimit;
  file print;
  put /// @5 'The Inputs:'
// @10 'Bias Estimate'                    @55 'Be'    @62 '=' @65 Be
  / @15 'The case is '                    @55 'Case'  @62 '=' @65 casenm
  / @15 'Standard Deviation of Bias Estimate' @55 'stdb'  @62 '=' @65 stdb
  / @15 'Degrees of Freedom of Bias Estimate' @55 'dfb'   @62 '=' @65 dfb1
  / @15 'Mean of logs of Study method obs.  ' @55 'sbar'  @62 '=' @65 sbar
  / @15 'Mean of logs of Indep method obs   ' @55 'ibar'  @62 '=' @65 ibar
  / @15 'Stdis from Equation 23             ' @55 'stdis' @62 '=' @65 stdis
  / @15 'StdD from  Equation 26             ' @55 'stdD'  @62 '=' @65 stdd
// @10 'Precision Estimate'                  @55 'SRTe'   @62 '=' @65 SRTe
  / @15 'Degrees of Freedom of SRT Estimate'  @55 'dfSRT'  @62 '=' @65 dfSRT
// @10 'Total Sample Size of the Test Method' @55 'n   '  @62 '=' @65 n
 / @10 'The Pump Error'                     @55 'pump'  @62 '=' @65 pump
     /// @5 'The Outputs:'
//                      @40 '95% Confidence Interval for Bias'
// @10 ' 2.5% Confidence Statistic for Bias'     @55 'B025'  @62 '=' @65 B025
 / @10 '97.5% Confidence Statistic for Bias'     @55 'B975'  @62 '=' @65 B975
//                      @40 '95% Confidence Interval for Precision'
// @10 ' 2.5% Confidence Statistic for SRT'      @55 'SRT025' @62 '=' @65 SRT025
 / @10 '97.5% Confidence Statistic for SRT'      @55 'SRT975' @62 '=' @65 SRT975
// @10 'The Bonferroni Approach:'
 /                      @40 '90% Confidence Interval for Accuracy'
 // @15 ' 5% Confidence Statistic for Accuracy'   @55 'BON05' @62 '=' @65 BON05
  / @15 '95% Confidence Statistic for Accuracy'   @55 'BON95' @62 '=' @65 BON95
   / @17 'Fulfill the Accuracy Criterion ?'  @55 BONpass
// @10 'The Hyperbolic Approach:'
 /                      @40 '90% Confidence Interval for Accuracy'
 // @15 ' 5% Confidence Statistic for Accuracy'   @55 'HYP05' @62 '=' @65 HYP05
  / @15 '95% Confidence Statistic for Accuracy'   @55 'HYP95' @62 '=' @65 HYP95
   / @17 'Fulfill the Accuracy Criterion ?'  @55 HYPpass;
run;
```

ILLUSTRATION - Output from execution of Computer Algorithm

Compute 5% and 95% Confidence Statistic for Accuracy.
Input Bias estimate, STD and DF of Bias Estimate,
SRT estimate, DF of SRT Estimate, and the Total Sample Size.

Output Confidence Statistics for Bias, SRT, and Accuracy.
BON05,BON95-Bonferroni Estimates. HYP05,HYP95-Hyperbolic Estimates
09:00 Monday, September 20, 1993

EXAMPLE 1 - II.B.2 IS not USED. neither EQUATION 23 nor 26 APPLY

The Inputs:

Bias Estimate Be = 0.03
The case is Case = STANDARD
Standard Deviation of Bias Estimate stdb = 0.04
Degrees of Freedom of Bias Estimate dfb = 30
Mean of logs of Study method obs. sbar = .
Mean of logs of Indep method obs ibar = .
Stdis from Equation 23 stdis = .
StdD from Equation 26 stdD = .

Precision Estimate SRTe = 0.07
Degrees of Freedom of SRT Estimate dfSRT = 15

Total Sample Size of the Test Method n = 18
The Pump Error pump = 0.05

The Outputs:

95% Confidence Interval for Bias
2.5% Confidence Statistic for Bias B025 = -0.051690898
97.5% Confidence Statistic for Bias B975 = 0.1116908983

95% Confidence Interval for Precision
2.5% Confidence Statistic for SRT SRT025 = 0.071777171
97.5% Confidence Statistic for SRT SRT975 = 0.1201526284

The Bonferroni Approach:
90% Confidence Interval for Accuracy
5% Confidence Statistic for Accuracy BON05 = 0.1406860352
95% Confidence Statistic for Accuracy BON95 = 0.3319702148
Fulfill the Accuracy Criterion ? Uncertain

The Hyperbolic Approach:
90% Confidence Interval for Accuracy

5% Confidence Statistic for Accuracy HYP05 = 0.1419777262
95% Confidence Statistic for Accuracy HYP95 = 0.2584988005
Fulfill the Accuracy Criterion ? Uncertain

Compute 5% and 95% Confidence Statistics for Accuracy.
Input Bias estimate, STD and DF of Bias Estimate,
SRT estimate, DF of SRT Estimate, and the Total Sample Size.
Output Confidence Statistics for Bias, SRT, and Accuracy.
BON05,BON95-Bonferroni Estimates. HYP05,HYP95-Hyperbolic Estimates
09:00 Monday, September 20, 1993

EXAMPLE 2 - II.B.2.b IS USED. EQUATION 26 APPLIES

The Inputs:

Bias Estimate Be = -0.029166211
The case is Case = II.B.2 USED
Standard Deviation of Bias Estimate stdb = .
Degrees of Freedom of Bias Estimate dfb = 18
Mean of logs of Study method obs. sbar = 6
Mean of logs of Indep method obs ibar = 6.0296
Stdis from Equation 23 stdis = .
StdD from Equation 26 stdD = 0.2

Precision Estimate SRTe = 0.07
Degrees of Freedom of SRT Estimate dfSRT = 15

Total Sample Size of the Test Method n = 18
The Pump Error pump = 0.05

The Outputs:

95% Confidence Interval for Bias

2.5% Confidence Statistic for Bias B025 = -0.120708153
97.5% Confidence Statistic for Bias B975 = 0.0719060456

95% Confidence Interval for Precision

2.5% Confidence Statistic for SRT SRT025 = 0.071777171
97.5% Confidence Statistic for SRT SRT975 = 0.1201526284

The Bonferroni Approach:
90% Confidence Interval for Accuracy

5% Confidence Statistic for Accuracy BON05 = 0.1406860352
95% Confidence Statistic for Accuracy BON95 = 0.3514404297
Fulfill the Accuracy Criterion ? Uncertain

The Hyperbolic Approach:
90% Confidence Interval for Accuracy
5% Confidence Statistic for Accuracy HYP05 = 0.1343015856
95% Confidence Statistic for Accuracy HYP95 = 0.2443959102
Fulfill the Accuracy Criterion ? YES

Compute 5% and 95% Confidence Statistics for Accuracy.
Input Bias estimate, STD and DF of Bias Estimate,
SRT estimate, DF of SRT Estimate, and the Total Sample Size.
Output Confidence Statistics for Bias, SRT, and Accuracy.
BON05,BON95-Bonferroni Estimates. HYP05,HYP95-Hyperbolic Estimates
09:00 Monday, September 20, 1993

EXAMPLE 3 - II.B.2.a IS USED. EQUATION 23 APPLIES

The Inputs:

Bias Estimate Be = -0.197481202
The case is Case = II.B.2 USED
Standard Deviation of Bias Estimate stdb = .
Degrees of Freedom of Bias Estimate dfb = 18
Mean of logs of Study method obs. sbar = 7
Mean of logs of Indep method obs ibar = 7.22
Stdis from Equation 23 stdis = 0.055
StdD from Equation 26 stdD = .

Precision Estimate SRTe = 0.0995
Degrees of Freedom of SRT Estimate dfSRT = 15

Total Sample Size of the Test Method n = 18
The Pump Error pump = 0.05

The Outputs:

95% Confidence Interval for Bias

2.5% Confidence Statistic for Bias B025 = -0.285055751
97.5% Confidence Statistic for Bias B975 = -0.099179521

95% Confidence Interval for Precision

2.5% Confidence Statistic for SRT SRT025 = 0.0885805689
97.5% Confidence Statistic for SRT SRT975 = 0.1634915275

The Bonferroni Approach:
90% Confidence Interval for Accuracy

5% Confidence Statistic for Accuracy BON05 = 0.2304382324
95% Confidence Statistic for Accuracy BON95 = 0.6661376953
Fulfill the Accuracy Criterion ? Uncertain

The Hyperbolic Approach:
90% Confidence Interval for Accuracy
5% Confidence Statistic for Accuracy HYP05 = 0.2824233932
95% Confidence Statistic for Accuracy HYP95 = 0.4340291137
Fulfill the Accuracy Criterion ? NO

Backup Data Report

Substance:	Cyclohexanone, No. S19
OSHA Standard:	200 mg/cu m
Chemical used for validation:	Cyclohexanone, Baker A.R.

General Procedure

The procedure followed for validation of the method for collecting and analyzing concentrations of cyclohexanone in air is described in NIOSH Method S19, which has been adapted from P&CAM 127. Desorption efficiency tests were done at 0.5, 1 and 2 times the OSHA standard by the method described in S19. Samples of cyclohexanone in dry air were generated by the procedure described in the Backup Data Report for cumene (Reference No, 1), and these samples were collected on activated coconut charcoal, Lot 105, supplied by SKC, Inc., Pittsburgh, Pa. Samples were collected for 40 minutes at a rate of approx. 1 liter per minute (individual critical orifices vary slightly in flow rate), The desorbed samples were analyzed by gas chromatography, and the amount measured was corrected for desorption efficiency (D.E.) by use of the D.E. curve. The concentration of cyclohexanone in air found was determined by dividing the corrected mg found by the sampled volume (critical orifice flow rate for that sample X 40 minutes). The true value for the concentration of cyclohexanone generated was determined by comparison with a bag standard as described in Reference No, 1, using a total hydrocarbon analyzer for measurement.

Modification of P&CAM 127

This method worked satisfactorily for cyclohexanone at the sample sizes generated using a 40 liter sample. Tests for desorption were first conducted using amounts of cyclohexanone equivalent to a 7.5 liter sample at 0.5X, 1X and 2X the OSHA standard and desorption efficiencies did not meet the 0.75 criterion specified by the Project Officer for this amount of analyte. The following solvents were tried for desorption:

Solvent	Amount	Approx. D.E. at 1X Standard
Carbon disulfide	0.5 ml	0.5
1% Methanol in carbon disulfide	0.5 ml	0.5
5% Methanol in carbon disulfide	0.5 ml	0.5
Methylene chloride	1.0 ml	0.25
2% Methanol in methylene chloride	1.0 ml	0.25

Methanol and ethanol were eliminated as possible desorbents since they interfered with gas chromatography of cyclohexanone.

Breakthrough Tests

A test for breakthrough of the front section of the charcoal tubes was conducted as described in the Backup Data Report for 2-butanone. (Reference 2.) Breakthrough occurred in 69 minutes when a concentration of 392 mg/cu m was sampled at a rate of 0.94 liter per minute.

Desorption Efficiencies

The analytical method was validated by performing desorption efficiency tests as described in the method. Results are given in the laboratory data section of this report.

References

1. Backup Data Report, Cumene, No. S23, prepared under NIOSH Contract CDC-99-74-45.

2. Backup Data Report, 2-Butanone, No. S3, prepared under NIOSH Contract CDC-99-74-45.

Laboratory Results

CYCLOHEXANONE

Analytical - Desorption Efficiencies (D.E.)*

0.5 OS			1.0 OS			2.0 OS		
mg Taken	mg Found	D.E.	mg Taken	mg Found	D.E.	mg Taken	mg Found	D.E.
3.764	2.894	0.769	7.53	6.02	0.800	18.82	16.12	0,855
3.764	2.988	0.794	7.53	6.37	0.846	18.82	16.68	0.886
3.764	2.746	0.730	7.53	5.94	0.789	18.82	17.01	0.904
3.764	2.748	0.730	7.53	6.14	0.816	18.82	17.30	O.919
3.764	2.723	0.723	7.53	6.12	0.814	18.82	16.75	0.890
3.764	2.848	0.7~7	7.53	6.02	0.800	18.82	17.54	0.932
Mean**		0.751			0.811			0.898
S***		0.027			0.020			0.027

Sampling and Analysis (values in mg/cu m)

Taken	Found	Taken	Found	Taken	Found
98.3	91.8	195.8	183.5	392	405
98.3	88.6	195.8	191.3	392	411
98.3	91.3	195.8	185.0	392	373
98.3	94.1	195.8	189.4	392	396
98.3	86.0	195.8	177.8	392	405
98.3	87.4	195.8	171.5	392	377
Mean**	89.9		183.1		395
S***	3.04		7.39		15.g
Error (%)****	-8.5		-6.5		+0.8

*D.E. = mg found/mg taken **Mean - Sum (D.E.)/n or Sum (Found)/n

***S = Standard Deviation ****Error (%) = 100(Found - Taken)/Taken

CYCLOHEXANONE

Statistical Analysis

Regression Equations:

Analysis:

$$Y = 0.804 - 0.280\ X_{1i} + 0.00579\ X_{2i}$$

t-values -1.717 2.586

Sampling and Analysis:

$$Y = 0.840 - 3.37\ X_{1i} + 0.00041\ X_{2i}$$

t-values 0.462 2.146

Coefficients of Variation:

CV_1^2	=	0.00063	Analysis
CV_2^2	=	0.00131	Sampling and Analysis
CV_P^2	=	0.0025	Pump Error
CV_T	=	0.0617	

where

$$CV_T^2 = CV_2^2 + CV_P^2$$

FAILURE REPORT

Substance:	Fluorotrichloromethane, No. S102
OSHA Standard:	1000 ppm (5600 mg/cu m)
Chemical used for validation:	Trlchlorofluoromethane, Certified Gas Mixture Matheson Gas Products

Discussion

This validation study was carried out using the small charcoal tubes. Although the CVT and recovery values are satisfactory, the small tubes ' capacity limits collection time to less than the required 50 minutes, even when sampling at 50 ml/min. For this reason, the method is considered a failure. Assuming that absorption capacity is proportional to the weight of charcoal in the tubes, use of the large charcoal tubes (ll cm long, 8 mm O.D. containing 400 mg of charcoal in the front section, 200 mg in the backup section) should provide adequate capacity (100 mg, 8 liters at 2X) to permit collection periods of 50 minutes (28 mg, 2.5 liters required). The respective breakthrough time for small and large tubes (front section only) for methyl chloride were ~2 and 10 minutes at 200 mg/min; for dichlorodifluoromethane the respective breakthrough times were 16 and 75 minutes at 50 ml/min. These data indicate that the capacity of the large tubes is at least four times as large as that of the small tubes.

Procedure

The procedure followed for validation of the method for collecting and analyzing concentrations of fluorotrichloromethane in air has been adapted from P&CAM 127. For determination of desorption efficiency, 100 mg samples of charcoal were placed in 2-ml Varian Automatic Sample Injector Vials and spiked with neat trlchloromonofluoromethane using a 10-µl syringe. The amount added was 7.3, 3.7, and 1.8 µl, respectively for the 2, l, and 0.5 times the OSHA standard level. This was the amount present in a l-liter air sample at the respective level. The spiking was done in a cold room because difluorodibromomethane boils at room temperature (B.P. = 23°C, d° = 1.5.). The eluting solvent was carbon disulfide.

For standards, a stock solution of fluorotrifluoromethane in carbon disulfide was prepared by adding 91 µl of neat monofluorotrichloromethane (from a 100 µl syringe) to 50 ml of carbon disulfide in a cold room. Dilutions were made at room temperature.

Samples of fluorotrichloromethane in air were generated and collected on activated coconut charcoal , Lot 105 , supplied by SKC , Inc ., Pittsburgh, Pennsylvania. The desorbed samples were analyzed by gas chromatography, and the amount measured was corrected for desorption efficiency D.E.) . The "found"' concentrations of fluorotrichloromethane in air were determined by dividing the corrected mg found by the sample volume (critical orifice flow rate for that sampLe X 20 minutes). The " true" concentrations of fluorotrichloromethane in the generated samples were determined by gas chromatographic analysis using a 5-ml sampling loop. The analysis was standardized by comparison with "bag" samples . The bag samples were prepared in 4-liter Teflon bags by metering 4 liters of nitrogen into the bag, followed by injecting the required amount of fluorotrichloromethane. Details of the analytical procedures are given in the Analytical Procedures section of Reference 1.

Modifications

The P&CAM 127 method was used with one modification. Two small charcoal tubes in series were used to avoid the problem of migration or fluorotrichloromethane from the sample section to the backup section after collection. The tubes were separated and capped immediately after sampling.

Generation

A gas mixture of fluorotrichloromethane in nitrogen at twice the OSHA standard level was purchased from Matheson Gas Products. The concentration was verified by gas chromatographic analysis using bag samples for comparison. These were prepared in 4-liter Teflon bags as described above. The test atmosphere was generated by directly sampling the mixture . The 3 concentrations were obtained by sampling only at the first stage and repeating the generation with appropriate dilution with nitrogen for concentrations 0.5 and 1 times the OSHA standard level. Complete mixing was assured by passing the stream of analyte through 2 mixers before sampling. The total gas flow rate was 2, 4, and 8 L/min for 0.5, 1, and 2 times the OSHA standard respectively. Sampling time was 20 minutes with a nominal flow rate of 50 mL/min for each charcoal tube. The details of the atmosphere generation equipment and operations are presented in the Atmosphere Generation Section of Reference 2.

Breakthrough

In order to test the capacity of the charcoal tubes an experiment was conducted at 2 times the OSHA standard level (actual concentration was 12502 mg/cu m). Breakthrough is defined as the time at which the effluent concentration from the tube reaches 5 percent of the concentration in the test gas mixture . The volume of sample to be used must be such that the volume of test air sampled at the time of breakthrough is greater than 1.5 times the volume of sample to be collected for analysis . In this breakthrough experiment, 3 tubes containing one section of 1OO mg of charcoal were used to sample the test air, which was pumped through the 8 tubes simultaneously through individual critical orifices. The combined effluent from the tubes was monitored continuously to detect breakthrough. Breakthrough occurred at 40 minutes. Thus the capacity of this lot of charcoal is 25 mg of fluorotrichloromethane and at the recommended sampling rate, the breakthrough volume is 2 liters.

Precision and Accuracy

The statistical procedures used are described in Reference 3 .

$CV_1 = 0.035 \quad CV_2 = 0.060 \quad CV_T = 0.079$

The average recovery or the generated samples over all levels was:

105.6%

References

1. Backup Data Report for Camphor, No. S10, NIOSH Contract No.CDC-99-74-45 .

2. Backup Data Report for Ethyl Alcohol, No. S56, ibid.

3 . Documentation of NIOSH Validation Tests, ibid.

4. Sarkan, A. E., and Greenburg, B. G., Contributions to Order Statistics, New York, John Wiley & Sons, 1962, p. 302.

DATA SHEET: FLUOROTRLCHLOROMETHANE S102

Analysis

Level 0. 5S			1S			2S		
mg Added	mg Found	DE	mg Added	mg Found	DE	mg Added	mg Found	DE
2.94	3.00	1.021	5.88	6.05	1.028	11.76	11.19	0.952
2.94	3.12	1.063	5.88	6.33	1.077	11.76	11.68	O.993
2.94	3.17	1.044	5.88	6.07	1.033	11.76	11.48	0.976
2.94	3.19	1.084	5.88	5.88	1.000	11.76	11.53	0.981
2.94	3.38	l.150	5.88	6.42	1.092	11.76	12.09	1.028
2.94	3.03	1.031	5.88	6.0	1.034	11.76	11.51	0.979
n =		6			6			6
mean		1.065			1.044			0.985
std dev		0 0470			0.0340			0.0251
CV_1		0.0441			0.0326			0.0254

CV_1 0 0349

CV_{A+DE} 0.0377

DATA SHEET: FLUOROTRICHLOROMETHANE S102

Sampling and Analysis

Test Level	Found mg	Corr mg	Liters	mg/cu m	Taken mg/cu m	Recovery
0.5S	3.036	2.843	0.906	3140	3050	102.9
	2.728	2.549	0.927	2751	3050	90.2
	3.292	3.089	0.909	3398	3050	111.4
	3.086	2.891	0.902	3204	3050	105.
	3.280	3.078	0.949	3245	3050	106.4
	1.846	1.717	0.913	1881*	3050	61.7*
		n = 5				
		mean		3147.60		103.2
		std dev		241.20		
		CV_2		0.077		
1S	6.957	6.749	0.906	7452	6285	118.6
	6.315	6.089	0.927	6571	6285	104.6
	6.282	6.055	0.909	6661	6285	106.0
	6.413	6.199	0.902	6860	6285	109.1
	6.487	6.265	0.949	6605	6285	105.1
	5.988	5.756	0.913	6306	6285	100.3
		n = 6				
		mean		6742.654		107.3
		std dev		390.471		
		CV_2		0.0579		
2S	11.824	12.008	0.906	13260	12502	106.1
	11.682	11.848	0.927	12787	12502	102.3
	11.938	12.137	0.909	13352	12502	106.8
	12.022	12.232	0.902	13558	12502	108.4
	12.466	12.737	0.949	13427	12502	107.4
	11.976	12.181	0.913	13344	12502	106.7
		n = 6				
		mean		13288.067		106.3
		std dev		265.128		
		CV_2		0.0199	CV_2	0.60

* Deleted as an outlier, because this value did not pass the Grubb's outlier test at the 1% confidence level as described in Reference No. 4.

FORMULA: Table 1 MW: Table 1 CAS: Table 1 RTECS: Table 1

METHOD: 1500, Issue 2 **EVALUATION: FULL** **Issue 1: 15 February 1984**

OSHA : Table 2
NIOSH: Table 2
ACGIH: Table 2

PROPERTIES: Table 1

COMPOUNDS: benzene, cyclohexane, n-heptane, n-hexane, n-octane, n-pentane
(Synonyms

SAMPLING

SAMPLER: SOLID SORBENT TUBE (coconut shell charcoal, 100 mg/50 mg)

FLOW RATE, VOLUME: Table 3

SHIPMENT: routine

SAMPLE STABILITY: at least 2 weeks

BLANKS: 2 to 10 field blanks per set

BULK SAMPLE: desirable, 1 to 10 mL; ship in separate containers from samples

ACCURACY

RANGE STUDIED: Table 3

BIAS: Table 3

OVERALL PRECISION ($\hat{S}_{rT}$): Table 3

ACCURACY: Table 3

MEASUREMENT

TECHNIQUE: GAS CHROMATOGRAPHY, FID

ANALYTE: hydrocarbons listed above

DESORPTION: 1 mL CS_2; stand 30 min

INJECTION VOLUME: 5 µL

TEMPERATURE-INJECTION: 250 °C
-DETECTOR: 250 °C
-COLUMN: see step 11

CARRIER GAS: N_2 or He, 25 mL/min

COLUMN: glass, 3.0 m x 2-mm, 20% SP-2100 on 80/100 mesh Supelcoport

CALIBRATION: analytes in CS_2

RANGE AND PRECISION: Table 4

ESTIMATED LOD: 0.001 to 0.01 mg per sample with capillary column [1]

APPLICABILITY: This method is intended for determining the OSHA-regulated hydrocarbons included within the boiling point range of n-pentane through n-octane. It may be used for simultaneous measurements; however, interactions between analytes may reduce breakthrough volumes and change desorption efficiencies.

INTERFERENCES: At high humidity, breakthrough volumes may be reduced by as much as 50%. Other volatile organic solvents, e.g., alcohols, ketones, ethers, and halogenated hydrocarbons, are likely interferences. If interference is suspected, use a more polar column or change column temperature.

OTHER METHODS: This method is based on and supercedes Methods P&CAM 127, benzene and toluene [2]; S28, cyclohexane [3]; S82, cyclohexene [3]; S89, heptane [3]; S90, hexane [3]; S94, methylcyclohexane [3]; S311, benzene [4]; S343, toluene [4]; S378, octane [4]; and S379, pentane [4]. For benzene or toluene in complex mixture of alkanes ($\leq C_{10}$), Method 1501 (aromatic hydrocarbons) is more selective.

REAGENTS:

1. Eluent: Carbon disulfide*, chromatographic quality with (optional) suitable internal standard.
2. Analytes, reagent grade.*
3. Nitrogen or helium, purified.
4. Hydrogen, prepurified.
5. Air, filtered.

* See SPECIAL PRECAUTIONS.

EQUIPMENT:

1. Sampler: glass tube, 7 cm long, 6-mm OD, 4-mm ID, flame-sealed ends, containing two sections of activated (600 °C) coconut shell charcoal (front = 100 mg, back = 50 mg) separated by a 2-mm urethane foam plug. A silylated glass wool plug precedes the front section, and a 3-mm urethane foam plug follows the back section. Pressure drop across the tube at 1 L/min airflow must be less than 3.4 kPa. Tubes are commercially available.
2. Personal sampling pump, 0.01 to 0.2 L/min, with flexible connecting tubing.
3. Gas chromatograph, FID, integrator and column (page 1500-1).
4. Vials, glass, 1-mL, with PTFE-lined caps.
5. Pipet, 1-mL, with pipet bulb.
6. Syringes, 5-, 10-, 25- and 100-µL.
7. Volumetric flasks, 10-mL

SPECIAL PRECAUTIONS: Carbon disulfide is toxic and extremely flammable (flash point = − 30 °C); benzene is a suspect carcinogen. Prepare samples and standards in a well-ventilated hood.

SAMPLING:

1. Calibrate each personal sampling pump with a representative sampler in line.
2. Break the ends of the sampler immediately before sampling. Attach sampler to personal sampling pump with flexible tubing.
3. Sample at an accurately known flow rate between 0.01 and 0.2 L/min (0.01 to 0.05 L/min for n-pentane) for a total sample size as shown in Table 3.
4. Cap the samplers with plastic (not rubber) caps and pack securely for shipment.

SAMPLE PREPARATION:

5. Place the front and back sorbent sections of the sampler tube in separate vials. Discard the glass wool and foam plugs.
6. Add 1.0 mL eluent to each vial. Attach crimp cap to each vial immediately.
7. Allow to stand at least 30 min with occasional agitation.

CALIBRATION AND QUALITY CONTROL:

8. Calibrate daily with at least six working standards over the appropriate range (ca. 0.01 to 10 mg analyte per sample; see Table 4).
   a. Add known amounts of analyte to eluent in 10-mL volumetric flasks and dilute to the mark.
   b. Analyze together with samples and blanks (steps 11, 12 and 13).
   c. Prepare calibration graph (peak area of analyte vs. mg analyte per sample).
9. Determine desorption efficiency (DE) at least once for each batch of charcoal used for sampling in the calibration range (step 8). Prepare three tubes at each of five levels plus three media blanks.

a. Remove and discard back sorbent section of a media blank sampler.
b. Inject a known amount of analyte directly onto front sorbent section with a microliter syringe.
c. Cap the tube. Allow to stand overnight.
d. Desorb (steps 5 through 7) and analyze together with working standards (steps 11, 12 and 13).
e. Prepare a graph of DE vs. mg analyte recovered.

10. Analyze three quality control blind spikes and three analyst spikes to insure that the calibration graph and DE graph are in control. Check for possible contamination during shipment of field samples by comparing results from field blanks and media blanks.

MEASUREMENT:

11. Set gas chromatograph according to manufacturer's recommendations and to conditions given on page 1500-1. Select appropriate column temperature:

	Approximate Retention Time (min), at Indicated Column Temperature			
Substance	40 °C	70 °C	100 °C	Programmed[a]
n-pentane	2.2	1.2		1.8
solvent (CS_2)	3.0	1.6		2.4
n-hexane	5.1	2.2		3.5
benzene[b]	7.7	3.2		4.5
cyclohexane[b]	8.4	3.4		4.7
cyclohexene	9.5	3.8		4.9
n-heptane	12	4.3		5.4
methylcyclohexane	14	5.2	2.2	5.9
toluene	17	6.5	2.6	6.5
n-octane	19	8.7	3.2	7.1

[a] Temperature program: 50 °C for 2 min, then 15 °C/min to 150 °C, 2-min final hold.
[b] Not completely resolved.

NOTE: Alternatively, column and temperature may be taken from Table 4.

12. Inject sample aliquot manually using solvent flush technique or with autosampler.
NOTE: If peak area is above the linear range of the working standards, dilute with eluent, reanalyze and apply the appropriate dilution factor in calculations.
13. Measure peak area.

CALCULATIONS:

14. Determine the mass, mg (corrected for DE) of analyte found in the sample front (W_f) and back (W_b) sorbent sections, and in the average media blank front (B_f) and back (B_b) sorbent sections.
NOTE: If $W_b > W_f/10$, report breakthrough and possible sample loss.
15. Calculate concentration, C, of analyte in the air volume sampled, V (L):

$$C = \frac{(W_f + W_b - B_f - B_b) \cdot 10^3}{V}, \ mg/m^3.$$

EVALUATION OF METHOD:

Precisions and biases (Table 3) were determined by analyzing generated atmospheres containing one-half, one, and two times the OSHA standard. Generated concentrations were independently verified. Breakthrough capacities were determined in dry air. Storage stability was not assessed. Measurement precisions (Table 4) were determined by spiking sampling media with amounts corresponding to one-half, one, and two times the OSHA standard for nominal air volumes. Desorption efficiencies for spiked samplers containing only one compound exceeded 75%. Reference [12] provides more specific information.

REFERENCES:

[1] User check, UBTL, NIOSH Sequence #4213-L (unpublished, January 31, 1984).
[2] NIOSH Manual of Analytical Methods, 2nd. ed., V. 1, P&CAM 127, U.S. Department of Health, Education, and Welfare, Publ. (NIOSH) 77-157-A (1977).
[3] NIOSH Manual of Analytical Methods, 2nd. ed., V. 2, S28, S82, S89, S90, S94, U.S. Department of Health, Education, and Welfare, Publ. (NIOSH) 77-157-B (1977).
[4] NIOSH Manual of Analytical Methods, 2nd. ed., V. 3., S311, S343, S378, S379, U.S. Department of Health, Education, and Welfare, Publ. (NIOSH) 77-157-C (1977).
[5] R. D. Driesbach, "Physical Properties of Chemical Compounds"; Advances in Chemistry Series, No. 15; American Chemical Society, Washington (1955).
[6] R. D. Driesbach, "Physical Properties of Chemical Compounds - II"; Advances in Chemistry Series, No. 22; American Chemical Society, Washington (1959).
[7] Code of Federal Regulations; Title 29 (Labor), Parts 1900 to 1910; U.S. Government Printing Office, Washington, (1989); 29 CFR 1910.1000.
[8] NIOSH Recommendations for Occupational Safety and Health. U.S. Department of Health and Human Services, DHHS (NIOSH) Publication No. 92-100 (1992).
[9] 1993 Threshold Limit Values for Chemical Substances and Physical Agents and Biological Exposure Indices. ACGIH, Cincinnati, OH (1993).
[10] Documentation of the NIOSH Validation Tests, S28, S82, S89, S90, S94, S311, S343, S378, S379, U.S. Department of Health, Education, and Welfare, Publ. (NIOSH) 77-185 (1977).

METHOD REVISED BY:

R. Alan Lunsford, Ph.D., based on results of NIOSH Contract CDC-99-74-45.

TABLE 1. SYNONYMS, FORMULA, MOLECULAR WEIGHT, PROPERTIES.

Name Synonyms	CAS# RTECS	Empirical Formula	Molecular Weight	Boiling Point (°C)	Vapor Pressure @ 25 °C (mm Hg)	Vapor Pressure @ 25 °C (kPa)	Density @ 20 °C (g/mL)
benzene[a] benzol; cyclohexatriene	71-42-2 CY1400000	C_6H_6	78.11	80.1	95.2	12.7	0.879
cyclohexane[a] hexahydrobenzene hexamethylene	110-82-7 GU6300000	C_6H_{12}	84.16	80.7	97.6	13.0	0.779
cyclohexene[a] tetrahydrobenzene benzene tetrahydride	110-83-8 GW2500000	C_6H_{10}	82.15	83.0	88.8	11.8	0.811
n-heptane[b]	142-82-5 MI7700000	C_7H_{16}	100.21	98.4	45.8	6.1	0.684
n-hexane[b] hexyl-hydride	110-54-3 MN9275000	C_6H_{14}	86.18	68.7	151.3	20.2	0.659
methylcyclohexane[a] cyclohexylmethane	108-87-2 GV6125000	C_7H_{14}	98.19	100.9	46.3	6.2	0.769
n-octane[b]	111-65-9 RG8400000	C_8H_{18}	114.23	125.7	14.0	1.9	0.703
n-pentane[b]	109-66-0 RZ9450000	C_5H_{12}	72.15	36.1	512.5	68.3	0.626
toluene[a] methylbenzene; toluol	108-88-3 XS5250000	C_7H_8	92.14	110.6	28.4	3.8	0.867

[a] Properties from [5].
[b] Properties from [6].

TABLE 2. EXPOSURE LIMITS, PPM [7-9].

Substance	mg/m³ OSHA TWA	OSHA C @ NTP		NIOSH Peak	NIOSH TWA	ACGIH C	ACGIH TLV	per ppm STEL
benzene*	10	25	50[b]	0.1[d]	1	10[d]		3.19
cyclohexane	300			300		300		3.44
cyclohexene	300			300		300		3.36
n-heptane	500			85	440	400	500	4.10
n-hexane[a]	500			50		50		3.52
methylcyclohexane	500			400		400		4.01
n-octane	500			75	385	300	375	4.67
n-pentane	1000			120	610	600	750	2.95
toluene	200	300	500[b]	100	150[c]	100	150	3.77

[a] The ACGIH recommendation for other hexane isomers is: TLV 500, STEL 1000.
[b] Maximum duration 10 min in 8 h.
[c] STEL
[d] Suspect carcinogen

TABLE 3. SAMPLING FLOWRATE[a], VOLUME, CAPACITY, RANGE, OVERALL BIAS AND PRECISION [2-4, 10].

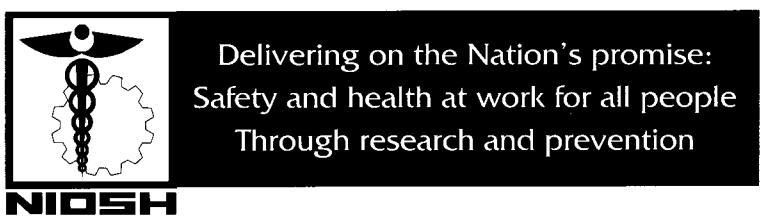
Substance	Sampling Flowrate (L/min)	Sampling Volume (L) MIN	Sampling Volume (L) MAX[b]	Breakthrough Volume at Concentration (L)	Breakthrough Volume at Concentration (mg/m^3)	Range at VOL-NOM (mg/m^3)	Overall Bias (%)	Overall Precision ($\hat{S}_{rT}$)	Accuracy (%)
benzene	≤0.20	2[c]	30	>45	149.1	41.5-165	0.4	0.059	±11.4
cyclohexane	≤0.20	2.5	5	7.6	1650	510-2010	1.1	0.060[d]	±11.5
cyclohexene	≤0.20	5	7	10.4	2002	510-2030	10.6	0.073	±20.7
n-heptane	≤0.20	4	4	6.1	4060	968-4060	-6.5	0.056	±15.0
n-hexane	≤0.20	4	4	5.9	3679	877-3679	-1.8	0.062	±12.5
methylcyclohexane	≤0.20	4	4	6.1	3941	940-3941	6.1	0.052	±15.2
n-octane	≤0.20	4	4	6.5	4612	1050-	-2.0	0.060	±12.1
n-pentane	≤0.05	2	2	3.1	5640	4403	-8.4	0.055	±16.6
toluene	≤0.20	2[c]	8	11.9	2294	1476-	1.6	0.052	±10.9
						6190			
						548-2190			

[a] Minimum recommended flow is 0.01 L/min.
[b] Approximately two-thirds the breakthrough volume.
[c] 10-min sample.
[d] Corrected value, calculated from data in [10].

TABLE 4. MEASUREMENT RANGE, PRECISION, AND CHROMATOGRAPHIC CONDITIONS [2-4,10].

	Measurement[a]		Carrier		Column Parameters[b]			
Substance	Range (mg)	Precision ($\hat{S}_r$)	Gas	Flow (mL/min)	t (°C)	Length (m)	Dia-meter (mm)	Packing[c]
benzene	0.09-0.35	0.036	N_2	50	115	0.9	3.2	A
cyclohexane	1.3 - 5.3[d]	0.024	N_2	50	210	1.2	6.4	B
cyclohexene	2.4 - 9.7[d]	0.021	N_2	50	205	1.2	6.4	B
n-heptane	4.08-16.3	0.016	He	30	80	3.0	3.2	C
n-hexane	3.56-14.5	0.014	He	30	52	6.1	3.2	D
methylcyclohexane	3.98-16.1	0.012	He	30	55	6.1	3.2	D
n-octane	4.75-18.9	0.009	He	30	52	6.1	3.2	D
n-pentane	2.98-11.8	0.014	He	30	52	6.1	3.2	D
toluene	1.13-4.51	0.011	N_2	50	155	0.9	3.2	B

[a] Injection volume, 5.0 µL; desorption volume, 1.0 mL, except cyclohexane and cyclohexene, 0.5 mL.
[b] All columns stainless steel. Diameter is outside dimension.
[c] A, 50/80 mesh Porapak P; B, 50/80 mesh Porapak Q; C, 10% OV-101 on 100/120 mesh Supelcoport; D, 10% FFAP on 80/100 mesh Chromosorb W AW-DMCS.
[d] Corrected value, calculated from data in [10].



DHHS (NIOSH) Publication No. 95-117