



Accuracy and Detection Limits for Bioassay Measurements in Radiation Protection

Statistical Considerations

**U.S. Nuclear Regulatory
Commission**

Office of Nuclear Regulatory Research

A. Brodsky



REPRODUCED BY
U.S. DEPARTMENT OF COMMERCE
NATIONAL TECHNICAL
INFORMATION SERVICE
SPRINGFIELD, VA 22161

NOTICE

Availability of Reference Materials Cited in NRC Publications

Most documents cited in NRC publications will be available from one of the following sources:

1. The NRC Public Document Room, 1717 H Street, N.W.
Washington, DC 20555
2. The Superintendent of Documents, U.S. Government Printing Office, Post Office Box 37082,
Washington, DC 20013-7082
3. The National Technical Information Service, Springfield, VA 22161

Although the listing that follows represents the majority of documents cited in NRC publications, it is not intended to be exhaustive.

Referenced documents available for inspection and copying for a fee from the NRC Public Document Room include NRC correspondence and internal NRC memoranda; NRC Office of Inspection and Enforcement bulletins, circulars, information notices, inspection and investigation notices; Licensee Event Reports; vendor reports and correspondence; Commission papers; and applicant and licensee documents and correspondence.

The following documents in the NUREG series are available for purchase from the GPO Sales Program: formal NRC staff and contractor reports, NRC-sponsored conference proceedings, and NRC booklets and brochures. Also available are Regulatory Guides, NRC regulations in the *Code of Federal Regulations*, and *Nuclear Regulatory Commission Issuances*.

Documents available from the National Technical Information Service include NUREG series reports and technical reports prepared by other federal agencies and reports prepared by the Atomic Energy Commission, forerunner agency to the Nuclear Regulatory Commission.

Documents available from public and special technical libraries include all open literature items, such as books, journal and periodical articles, and transactions. *Federal Register* notices, federal and state legislation, and congressional reports can usually be obtained from these libraries.

Documents such as theses, dissertations, foreign reports and translations, and non-NRC conference proceedings are available for purchase from the organization sponsoring the publication cited.

Single copies of NRC draft reports are available free, to the extent of supply, upon written request to the Division of Technical Information and Document Control, U.S. Nuclear Regulatory Commission, Washington, DC 20555.

Copies of industry codes and standards used in a substantive manner in the NRC regulatory process are maintained at the NRC Library, 7920 Norfolk Avenue, Bethesda, Maryland, and are available there for reference use by the public. Codes and standards are usually copyrighted and may be purchased from the originating organization or, if they are American National Standards, from the American National Standards Institute, 1430 Broadway, New York, NY 10018.

NRC FORM 335 (2-84) NRCM 1102 3201, 3202 SEE INSTRUCTIONS ON THE REVERSE	U.S. NUCLEAR REGULATORY COMMISSION BIBLIOGRAPHIC DATA SHEET	1 REPORT NUMBER (Assigned by TIDC, add Vol. No., if any) NUREG-1156
2 TITLE AND SUBTITLE Accuracy and Detection Limits for Bioassay Measurements in Radiation Protection - Statistical Considerations	3 LEAVE BLANK	
5 AUTHOR(S) A. Brodsky	4 DATE REPORT COMPLETED MONTH: February YEAR: 1986	
7 PERFORMING ORGANIZATION NAME AND MAILING ADDRESS (Include Zip Code) Division of Radiation Programs and Earth Sciences Office of Nuclear Regulatory Research U. S. Nuclear Regulatory Commission Washington, DC 20555	6 DATE REPORT ISSUED MONTH: April YEAR: 1986	
10. SPONSORING ORGANIZATION NAME AND MAILING ADDRESS (Include Zip Code) Same as 7 above.	8. PROJECT/TASK/WORK UNIT NUMBER OP 404-1	9 FIN OR GRANT NUMBER
12 SUPPLEMENTARY NOTES	11a TYPE OF REPORT Technical report b PERIOD COVERED (Inclusive dates)	
13 ABSTRACT (200 words or less) <p>This report presents statistical concepts and formulas for defining minimum detectable amount (MDA), bias and precision of sample analytical measurements of radioactivity for radiobioassay purposes. The defined statistical quantities and accuracy criteria were developed for use in standard performance criteria for radiobioassay, but are also useful in intralaboratory quality assurance programs. This report also includes a literature review and analysis of accuracy needs and accuracy recommendations of national and international scientific organizations for radiation or radioactivity measurements used for radiation protection purposes. Computer programs are also included for calculating the probabilities of passing or failing multiple analytical tests for different acceptable ranges of bias and precision.</p>		
14 DOCUMENT ANALYSIS - a KEYWORDS/DESCRIPTORS radiobioassay, bioassay, radiation protection, health physics, radiation measurements, quality assurance of radiation and radioactivity measurements, statistics of radiation measurements, accuracy of radiation measurements, detection limits, minimum detectable amount, bias, precision b IDENTIFIERS/OPEN-ENDED TERMS	15 AVAILABILITY STATEMENT Unlimited	16 SECURITY CLASSIFICATION (This page) Unclassified (This report) Unclassified
		17 NUMBER OF PAGES 145
		18 PRICE

Accuracy and Detection Limits for Bioassay Measurements in Radiation Protection

Statistical Considerations

Manuscript Completed: February 1986
Date Published: April 1986

A. Brodsky

**Division of Radiation Programs and Earth Sciences
Office of Nuclear Regulatory Research
U.S. Nuclear Regulatory Commission
Washington, D.C. 20555**



ABSTRACT

This report provides statistical concepts and formulas for defining minimum detectable amount (MDA), bias and precision of sample analytical measurements of radioactivity for radiobioassay purposes. The defined statistical quantities and accuracy criteria were developed for use in standard performance criteria for radiobioassay, but are also useful in intralaboratory quality assurance programs. This report also includes a literature review and analysis of accuracy needs and accuracy recommendations of national and international scientific organizations for radiation or radioactivity measurements used for radiation protection purposes. Computer programs are also included for calculating the probabilities of passing or failing multiple analytical tests for different acceptable ranges of bias and precision.



TABLE OF CONTENTS

<u>CHAPTER</u>	<u>Page</u>
ABSTRACT	ii
PREFACE	xi
1 INTRODUCTION.....	1
2 RATIONALE FOR METHODS OF DETERMINING MINIMUM DETECTABLE AMOUNT (MDA) AND ACCEPTABLE MINIMUM DETECTABLE AMOUNT (AMDA)	2
2.1 Reasons for Establishing Standard Definitions and Formulations of MDA and AMDA	2
2.2 Previous Formulations of MDA	3
2.3 Formulation of MDA for Radiobioassay Performance Standards	5
2.4 Important Points and Examples in the Use of the MDA Formulation	8
2.4.1 Need to Use Appropriate Values of Each Variable	8
2.4.2 The Standard Error of the Count Rate	8
2.4.3 Blank Signals with Gaussian Variations	9
2.4.4 Formulation of MDA for Variable Calibration or Conversion Factor K	9
2.4.5 Estimation of MDA in the Presence of Known Quantities of Interfering Nuclides	11
2.4.6 Currie's MDA (LLD) Formulation, Including Adjustments for Limited Systematic Errors in Blank and Calibration Factor	16
2.4.7 MDA Formulation for Rapidly Decaying Nuclides	21
2.4.8 MDA Determinations When the Physical Processes of Measurement are Stable and Well Known	22
2.4.9 MDA Formulations for Measurement Processes Having Continuous Output Signals	22
2.4.10 Measurements in Uncontrolled Environments	25
2.4.11 Multiple Detection Systems for More Than One Nuclide	25
2.4.12 Detection by Visual Observation	25
2.4.13 Examples of MDA Calculations	26
2.5 Recording of Analytical Results	28
2.6 Considerations in Selecting Representative MDA's	29
2.7 Considerations for Selection of Acceptable Minimum Detectable Amounts (AMDA's)	29

TABLE OF CONTENTS (Continued)

<u>CHAPTER</u>		<u>Page</u>
3	RATIONALE AND PRACTICAL IMPLICATIONS OF THE BIAS AND PRECISION STATISTICS, AND THE CRITERIA FOR TEST SAMPLE ACCURACY	33
3.1	General Approach	33
3.2	Demonstration of the Unbiasedness of B_{r_i}	34
3.3	The Unbiasedness of the Precision Measure S_B^2	36
3.4	Other Desired Characteristics of the Precision Statistic S_B	37
3.5	Characteristics of the Precision Statistic, S_A , the Coefficient of Variation	38
3.6	Rationale for the Limits on Bias and Precision	39
3.6.1	Specific Accuracy Criteria Proposed	39
3.6.2	General Considerations in Establishing Standards of Accuracy	40
3.6.3	Summary of the Literature and Considerations Regarding Accuracy Requirements	40
3.6.3.1	Radiation Protection Accuracy Requirements	40
3.6.3.2	Administrative Considerations	44
3.6.3.3	Accuracy Requirements for Bioassay as Implied by Uncertainties in Air Monitoring	46
3.6.3.4	Legal Requirements for Accuracy	51
3.6.3.5	Accuracy Requirements in Management of Emergencies Involving Radionuclide Intake	55
3.6.3.6	Accuracy Needs for Epidemiologic Investigations of Radiation Risk	57
3.6.4	Conclusions Regarding Accuracy Requirements of Bioassay, Specifications of Bias and Precision Statistics and Acceptable Ranges for Sample Quality Control Tests	58
3.6.4.1	General Conclusions	58
3.6.4.2	Specific Conclusions Regarding the Selection of Specific Accuracy Criteria and Practical Statistical Estimators	59
3.7	Implications of Finite Sample Sizes in Terms of Failing the Precision Criteria	60
3.8	A Computer Program for Calculating the Probability of Passing the Precision Criteria for a Replicate Test Measurement of N Spiked Samples	61

TABLE OF CONTENTS (Continued)

<u>CHAPTER</u>	<u>Page</u>
3.9 Implications of Finite Sample Sizes in Terms of Passing the Bias Criteria	65
3.9.1 Probabilities to be Calculated	65
3.9.2 Probability Statements for Normally-Distributed Values of A_{aj}	65
3.9.3 Probability Statements for Lognormally Distributed Values of A_{aj}	66
3.9.4 Calculation of the Probability of \bar{B}_r Falling Within a Specified Range	68
3.10 Computer Program and Example Calculations	70
3.11 Example of Calculation of Probability of Passing Both the Bias and Precision Tests for $N = 3$	72
4 DISCUSSION	78
REFERENCES	79

TABLE OF CONTENTS (Continued)

LIST OF TABLES

<u>TABLE</u>	<u>Page</u>
1 Probabilities of Failing Precision Test versus Relative Error and Sample Size	61
2 True Analytical Precision Necessary to Pass Relative Precision Criteria $ S_B \leq 0.4$, For Various Sample Sizes	62
3 Calculations of the Values of 1-FCHISQ, the Probability of Failing the Precision Test in Table 1, Using the Computer Programs of Exhibit 1	64
4 Probabilities of Passing Bias and Precision Tests (Requirements of True Precision and Bias to Pass Multi-Category Tests With Only Two Pass/Fail Criteria of $ S_B $ or $ S_A \leq 0.4$ and $-0.25 \leq B_r \leq 0.50$, when these Statistics are Stochastically Independent)	74

TABLE OF CONTENTS (Continued)

LIST OF FIGURES

<u>FIGURE</u>		<u>Page</u>
1	Decision Level, L_C	30
2	Detection Limit, L_D (Paired Blank)	31
3	Zero Background (Blank) Count	32
4	Finding Parameters for Normal A_{ai}	76
5	Finding Parameters for Lognormal A_{ai}	77

TABLE OF CONTENTS (Continued)

LIST OF EXHIBITS

<u>EXHIBIT</u>		<u>Page</u>
1	Program for Calculating Chance of Passing Precision Tests and Subroutine for the Gamma Function	63
2	Program for Computing $P(b_1 \leq B_r \leq b_2)$ for Either Normally or Lognormally Distributed Sample Measurements	75

PREFACE

This report presents a review of statistical concepts and formulas for the minimum detectable amount (MDA) of an analytical determination, definitions of "sample" bias estimators for use in interlaboratory comparisons or accreditation programs, and criteria for evaluating the performance of laboratory analyses in terms of detection limits (acceptable MDA's) and acceptable levels of accuracy expressed in terms of bias and precision statistics. Also, a literature review and rationale are presented for choosing certain accuracy criteria, at least for initial use in standard criteria for radioanalytical determinations in the field of radiobioassay. The word "radiobioassay" in this report refers to the determination of quantities of radioactive materials in biologic media, including the human body in vivo and materials removed from the body in vitro.

The author first proposed the use of certain relative bias and precision estimators, combined with certain accuracy criteria, during early (1981-82) discussions as a member of Working Group 2.5 of the Health Physics Society Standards Committee. This Working Group was charged with the development of a draft standard ANSI N13.30, Performance Criteria for Radiobioassay. This Working Group was initiated by the Health Physics Society in 1981, upon receiving recommendations from Robert E. Alexander, U.S. Nuclear Regulatory Commission and Edward J. Vallario, Department of Energy.

However, the need for such a broad standard of accuracy and performance reliability for radiobioassay, both for intralaboratory quality control and interlaboratory performance testing and accreditation, had been increasingly recognized by radiobioassay chemists and health physicists over the last three decades. This recognition occurred in the private commercial laboratories as well as in governmental agencies concerned with monitoring or auditing the accuracy and reliability of measurements made to protect human health. Of more than 30 persons from private commercial laboratories offering radiobioassay services, about 75% of those surveyed at the 1982 Conference on Bioassay, Environmental and Analytical Chemistry indicated that they supported not only a consensus standard on this subject, but also a uniformly conducted national accreditation program, and an NRC regulation requiring licensees who relied on bioassay measurements for regulatory purposes to use accredited laboratories.

Because of the broad interest and support for this standard, and its potentially broad applicability to many laboratories performing analyses important to human health, both the Working Group and my supervisor, Mr. Robert E. Alexander, encouraged a very careful early development of the ANSI N13.30 draft. Kenneth Heid brought additional consultants and members into the Working Group as their expertise was needed, and successive drafts of the standard and the rationale for the approach to the quantitative statistical criteria were circulated to more than 100 persons expert in the radiochemical, radiometric, *in vivo* counting, mathematical modelling, health physics management, and statistical aspects of radiobioassay. Many of these persons are among the leading scientists in the world in fields related to radiobioassay. Many of the concepts in this report, as well as in the standard, were revised and polished several times utilizing suggestions and information provided by these experts. During the period 1982-85, the Working Group added additional criteria for the precision statistics,

and changes to the recommended formulations of MDA, as a result of comments received from the members of the Working Group and the other experts. Also, over two hundred references in the scientific literature were examined to check scientific data and opinion related to the criteria under development by the working group, and to establish a scientific basis and rationale for the quantitative statistical criteria for use in radiobioassay quality control and quality assurance.

While reviewing the literature, it became obvious to the author that there are many thousands of scientific articles related in some way to the subject of this report. This is not surprising since the statistical concepts of measurement processes applicable to radiobioassay are generally applicable to many other types of scientific measurements. Thus, the author has also relied on authoritative review articles and on direct consultations with authorities in the field of analytical measurements. Still, the author welcomes further comments on this report, since many persons whose review of this material was solicited did not have time for a thorough review and there are hundreds of other scientists who will have an interest and perhaps desire to suggest useful additions or changes to material in the draft standard. Indeed, one of the purposes of publishing this report at this time is to obtain a wider distribution and review of this material while the development and review of the bioassay performance criteria are still underway.

Another purpose of publication of this report is to make the detailed considerations and rationale for selecting certain quantitative performance criteria available to those who may be required to consider the necessity of future revisions after the final ANSI standard is published. The Working Group at the time of this writing (January 1986) is still in the final stages of preparing the draft standard, Performance Criteria for Radiobioassay.

Most of the material in this report was originally abstracted and consolidated from information in about 130 publications as well as from analytical and mathematical work carried out as part of the Working Group effort from 1981-1985. This material was originally written to be included in appendices to the ANSI standard. However, in its meetings in 1985, the Working Group decided to include only a summary of the material presented here in appendices to the standard, recommending that the detailed mathematical derivations and statistical considerations be published as a separate report. This approach was also encouraged by R. E. Alexander of the NRC. Thus, this report is published as an independent report of the author, for its possible usefulness in examining and understanding the bases for the specific quantitative accuracy and MDA formulations and criteria in the ANSI standard, and for its possible auxiliary and independent use by scientists involved in the quality assurance of analytical programs.

Although many have contributed to the ideas, approaches and methods presented in this report, a special acknowledgement must be made to the Working Group members, who include chemists, physicists, mathematicians and health physicists, and their employers, who gave generously of their time and energy for the development of radiobioassay criteria. These individuals will be listed in the ANSI N13.30 standard. Specific reviews at various stages that provided suggestions influential in developing the statistical approach for calculating MDA and bias and precision statistics were made by Dr. Charles T. Schmidt, University of California; Drs. Lloyd A. Currie, Kenneth W. G. Inn, J. M. R. Hutchinson and

Walter Liggett, National Bureau of Standards; Alfred V. Robinson and Mathias M. Lardy, U.S. Testing Company; Dr. David E. McCurdy and Russell A. Mellor of Yankee Atomic Electric Company; Dr. Darrell A. Fisher, Battelle Pacific Northwest Laboratories; Claude W. Sill, Idaho Falls; Roscoe M. Hall, Jr., DuPont de Nemours and Company, Savannah River Laboratory; Roger Falk, Rockwell International; William D. Moss, Los Alamos National Laboratory; John R. Buchanan, and Dr. Michael E. Ginevan, Nuclear Regulatory Commission; C. M. West of Oak Ridge, Tennessee and Joseph C. Lochamy of Clinton, Tennessee. Also, the author acknowledges the special encouragement and generally helpful direction provided by Kenneth R. Heid, Chairman of the ANSI N13.30 Working Group, Robert E. Alexander, Nuclear Regulatory Commission, who provided much of the initial incentive and programmatic support for the working group effort and the parallel interlaboratory testing project at Battelle Pacific Northwest Laboratories, and Edward J. Vallario Department of Energy, who also helped initiate the development of the ANSI standard and continues to provide the needed support for the research to adequately test the numerical criteria in the standard. The author also acknowledges the support of NRC management in making his efforts available to the Working Group, and to the projects at Battelle Pacific Northwest Laboratories (NRC FIN NO. B2417) and the National Bureau of Standards (NRC FIN No. B8093), which were necessary to provide much of the test data needed to develop the statistical formulations of test criteria and the quantitative accuracy criteria.

Although many have been helpful in the development of information necessary to prepare this report, any opinions or judgments in the way this information was synthesized into the material in this report are the sole responsibility of the author. The author hopes and expects that this report will be helpful to the ANSI N13.30 Working Group and future working groups that may be called upon to further develop or revise the ANSI N13.30 or similar standards. However, as an independent document, the statements in this report do not necessarily represent the opinions or policies of the Working Group, the Health Physics Society, or the author's employer, the U.S. Nuclear Regulatory Commission. Further, it is not intended or represented that this document deal with all possible approaches or viewpoints on the basic subject matter of quality control of analytical measurements. The author would appreciate receiving any further comments or suggestions on the subject matter of this document, to help ensure the optimum input of scientific expertise into the important efforts to improve the sensitivity, accuracy and reliability of analytical determinations of radioactivity in the human body or in samples used for the assessment of human intakes of radioactive material.



1 INTRODUCTION

The purpose of this report is to present the rationale and statistical formulations for determining the minimum detectable amount (MDA), bias and precision of an analytical determination for purposes of the radiobioassay of radioactive materials in the human body or in samples of biological material from the body. The formulations were developed for use in standardizing performance criteria for radiobioassay, and provide simple formulas from which the statistics desired can be computed from replicate sample determinations of concentrations or amounts of radioactive material in biological media. Acceptable levels of bias and precision of radiobioassay analyses are also discussed in terms of the sample bias and precision statistics, with a review of selected references from the literature, and references to statements of national and international scientific organizations regarding the degree of accuracy of radiation measurements and radioanalytical determinations for various radiation risk assessment and radiation protection purposes.

The rationale, statistical concepts reviewed, statistical estimator formulations and the methods for evaluating the implications of pass/fail criteria in laboratory intercomparison programs, as presented in this report, may be useful in individual laboratory quality control programs, as well as in present and future development of national standards of both intra- and inter-laboratory quality assurance. Algorithms are also provided for evaluating the probabilities of passing multiple bias and precision tests, with replicated sample measurements in multiple categories of analysis, for any given ranges of acceptable bias and precision in the sample estimators, versus given underlying true biases and precisions of analytical determination. These algorithms are derived for both Gaussian and lognormal random errors of measurement, and for biases in both additive and multiplicative factors in the radioanalytical determination.

2 RATIONALE FOR METHODS OF DETERMINING MINIMUM DETECTABLE AMOUNT (MDA) AND ACCEPTABLE MINIMUM DETECTABLE AMOUNT (AMDA)

2.1 Reasons for Establishing Standard Definitions and Formulations of MDA and AMDA

Clear formulations of the concepts of "minimum detectable amount (MDA)" and "acceptable MDA (AMDA)", and standard methods of determining them, are important not only for the purposes of developing standards for specifying bioassay laboratory performance criteria, but are also more broadly significant for the effective management and quality control of bioassay services and for the proper auditing of personnel exposures in radiation protection programs. Without a standard of practice, various laboratories have in the past defined either their "lower limits of detection (LLD)", or MDA's, in various terminologies and formulations (1). This situation has made it difficult to intercompare laboratory capabilities, and has had other detrimental influences on the science and commerce of bioassay laboratories.

The advantages of standardizing the concepts of MDA and AMDA can be itemized as follows:

1. Proper determination of the MDA of a measurement prevents false claims that a given low amount of radioactivity can be detected with a high degree of assurance. Although false claims may bring some initial business to a laboratory, the erroneous reporting of zero or negligible quantities when positive test "spiked" samples above the claimed MDA have been submitted to the laboratory will soon result in loss of reputation and business decline.
2. On the other hand, proper determination of MDA may avoid understatements of capability, and thus loss of information or business resulting from a belief that an analytical process has less detection capability than it actually does have.
3. A proper determination of MDA, which requires estimates or determinations of basic measurement variances, will allow proper projections of the accuracies of internal dose assessments from multiple bioassay samples. In the past, understatements of the detectability of personnel monitoring films (2), coupled with a lack of consideration of the power of the statistical (and experimental) laws of large numbers (3), have led to major biases in reporting and/or recording cumulative personnel doses due to improper dropping of significant digits, or on the other hand the practice where all doses are conservatively recorded as at least equal to some presumed and overestimated "detection limit". Of course, errors of bioassay measurements must be compounded with the additional uncertainties of physiological modeling and biological variability to determine overall uncertainties in internal dose estimates (see Section 3). Still, assessments of internal exposure are subject to many biases of reporting and recording, and unaccountable differences between laboratories, without standardized and properly applied methods of determining MDA's.
4. Proper determination of MDA and understanding of the magnitudes and sources of error in bioassay analyses can be important in establishing records presentable with confidence in court proceedings, where such records have

shown the need to stand up to close scrutiny and questioning (see Reference 4 and Section 3).

5. Determination and standardization of the AMDA for radiation protection purposes first requires the proper determination of MDA for those nuclides that are more difficult to measure at low levels.

It would be counter to radiation protection purposes to establish AMDA's lower than MDA's that are practicably attainable by competent analytical scientists using the best state-of-the-analytical-art at economically affordable costs to the customer. This would discourage the better laboratories from offering at all some of the most needed analyses for some of the most radiotoxic materials. Further, an implicit, if not explicit, cost-benefit consideration is necessary in selecting an AMDA sufficiently higher than the MDA so that not only is detection of a radionuclide assured for reasonably competent attention to procedural quality control, but also the required analyses can be made at costs and within times consistent with the need to allow ample numbers of sample measurements to adequately monitor the exposed populations of interest.

6. The standardization of the formulations of MDA and the selections of AMDA through an accepted industry standards-development process, involving input from a broad group of scientists and a wide segment of the industries concerned, can reduce confusion and chaos in the planning and advertising of laboratory services as well as in the selection of service laboratories by industries needing specific kinds of bioassay determinations.

Thus, the broad importance to industry involved with bioassay of radioactive materials of developing standard methods for determining MDA, and standards of AMDA, is clear. This section will summarize the rationale for the formulation of MDA for use in bioassay performance standards, and will introduce considerations for selecting particular values of AMDA for different nuclides and categories of measurement.

2.2. Previous Formulations of MDA

There have been a number of attempts to improve the conceptual and verbal consistency of defining the minimum amount of radioactive material that is detectable at given pre-selected probability levels of confidence (5-16). However, this appendix will focus primarily on those based on the analysis of Currie (7), whose terminology and derivations rely on familiar statistical and measurement concepts, and are applicable to the approach to be suggested here. Also, Currie's basic formulation of detection limit, L_D , in terms of a net count that is 4.65 times the standard deviation of an appropriate blank sample, has been widely adapted for use in defining MDA for practical measurements in situations where the total counts are high enough so that Poisson variations are adequately approximated by the Normal (Gaussian) distribution for purposes of establishing probability levels (11, 14-16). About 50 counts are needed in order to use Normal distribution tail probabilities of 0.01 or greater (17), but at the 0.05 level, the 4.65 S formulation has sometimes been used down to only a few counts (11).

The following equation for LLD is typical of those that have been adapted for use in recent years (9-12, 14-16) for describing the analytical capabilities of laboratory analyses of environmental samples of radioactivity (14):

$$LLD = \frac{4.65s_b}{E \sqrt{V \cdot 2.22 \times 10^6 Y \exp(-\lambda \Delta t)}} \quad \text{Eq (1)}$$

s_b = standard deviation of blank, counts/minute,

E = counting efficiency, counts/disintegration,

V = sample size in units of mass or volume,

2.22×10^6 = number of disintegrations/min per μCi ,

Y = fractional radiochemical yield, when applicable,

λ = radiological decay constant,

Δt = elapsed time between sample collection and counting

For the factors in the denominator, Reference 14 states that "Typical values should be used in the denominator." When chemical yields are variable and frequently low, this practice could result in an exaggeration of detection capabilities. It is therefore important in the use of the above formulation to be sure that it is applied only when a sufficient number of counts are obtained for the blank, or the measurement is for other reasons normally distributed, since Currie's (7) constant term of 2.71 is omitted. It is also important to provide quality control that ensures that the factors in the denominator do not vary too widely.

Although not applicable to situations where the blank counts are very low or zero, Equation 1 otherwise ensures about 95% chance or greater that an LLD quantity of radioactive material (or more) will be detected. These probabilities apply when a paired blank is counted (or measured) for the same total counting time as the sample. If in the calculation of a sample result, the subtracted blank value is taken to be a constant from a well-known, stable blank count, the 4.65 constant may be replaced by 3.29 (7). However, as long as the blank count is well above zero, the generic use of Equation 1 would ensure that the chance of detecting an LLD quantity or more would be at least 95% (and the probability of a Type II error would be less than 0.05).

The more recent paper by Donn and Wolke (13) suggests a method for defining a "mean probability of detection, Q_d ", which attempts to weight the integration of the probability of obtaining a count above the detection level by an assumed probability density (Normal) distribution of the underlying mean count μ given a count C . However, this concept is as arbitrary a definition as any that provides a consistent way of expressing LLD, or MDA, in terms of L_D . Donn and Wolke's methods may be useful where one wishes to examine the effects of choosing various probability levels, and also adjust the relative counting times of background and sample. However, their formulation is rather complex and employs the inherent assumption that sufficient counts are obtained for both the sample and the blank, so that the normal probability densities are applicable; this does not provide a formulation applicable to many present measurement systems

involving low-background spectroscopy. Further, corrections would be needed to the Donn-Wolke approach (18).

Also, the Currie formula as modified in the following section with a changed constant term would be more conservative in terms of assuring that detection probabilities are at least not lower than those deemed to be adequate a priori, and would be relatively accurate for both low and high background situations when background and calibration or conversion factors are stable. Guidance will be presented for calculating MDA's when calibration or conversion factors may have random or fixed errors.

In developing the following formulation of MDA, as in developing practical and efficient criteria for bias and precision in Section 3, an attempt has been made to use terminology that is at least consistent with recent statistical practices (19).

2.3. Formulation of MDA for Radiobioassay Performance Standards

A standard definition of MDA has been developed that can be applied to the various types of measurement systems used in bioassay, with the assurance that any claimed MDA would have at least a 95% chance of being detected. The rationale for the formulation of MDA is summarized below.

The decision level, L_C , and Detection Limit, L_D , definitions were utilized as given by Currie (7) (See Figures 1 and 2). The Determination Level concept (7) may be used in developing the activity levels at which various quality control test procedures should be conducted, for either interlaboratory or intralaboratory testing. As shown in Figure 2, L_D is the amount that would give an expected signal (or a net count over the total counting time interval) such that the probability of obtaining a net count less than L_C (and thus a "false negative" result) would be 0.05.

The equation for detection limit of a net count, when a sample is counted over the same time period as a paired blank, has been derived by Currie (7) as the following approximation under the assumption that Normal distribution probability intervals are applicable:

$$L_D = 2.71 + 4.65 s_b, \quad \text{Eq (2)}$$

where s_b (an estimate of σ_b) is the standard deviation in repeated measurements of an appropriate blank, when each blank is counted for the same counting time interval as the sample. The following equations used by Currie (7) in deriving Equation 2 may help provide a better understanding of the conditions under which Equation 2 is apt to provide a valid formulation:

Currie (7) established a "decision limit, L_C " by setting L_C at a distance (see Figure 1) from zero so that

$$L_C = k_\alpha (\mu_B + \sigma_B^2)^{\frac{1}{2}} \quad \text{Eq (3)}$$

In this way, the use of μ_B recognizes when few counts are obtained that the variance may significantly depend on μ_B . However, in deriving the formulation for L_C , Currie substitutes $\mu_B = \sigma_B^2$ and the Normal one-tailed 0.05 critical k_α into Eq. 3 to obtain:

$$L_C = 1.645 (2 \sigma_B^2)^{\frac{1}{2}} = 2.33 \sigma_B \quad \text{Eq (4)}$$

The formula for L_D was obtained (7) setting

$$L_D = L_C + k_\beta \sigma_D = L_C + k_\beta (L_D + \sigma_D^2)^{\frac{1}{2}}, \quad \text{Eq (5)}$$

which establishes the distance of L_D away from L_C so that the probability of a Type II error is β (see Figure 2).

Then, Equation 5 was solved for L_D , algebraically simplified, k_α and k_β were set equal to yield:

$$L_D = k^2 + 2L_C \quad \text{Eq (6)}$$

which with the substitution of $k = 1.645$ for the one-tailed Normal distribution probability level of 0.05, and the substitution for L_C from Equation 4, becomes

$$\begin{aligned} L_D &= (1.645)^2 + 2(2.33) \sigma_B \\ &= 2.71 + 4.65 \sigma_B \end{aligned} \quad \text{Eq (7)}$$

which is the same as the Equation 2 as derived by Currie (7).

This brief summary of Currie's derivation shows clearly that the constant 2.71 is an artifact of the mixed assumptions of Normal and Poisson-distributed variability of measurements. As discussed by Currie (7), the constant 2.71 does not provide an 0.05 probability of a Type II error in the limit of very small or zero values of σ_B (i.e., very small or zero total blank counts).

Figures 1 and 2 picture the approximate probability distributional situations when the 4.65 s term is dominant and the 2.71 term may be neglected. The distribution to the right is wider due to the assumption $\sigma^2 = \mu$. Figure 3 pictures the situation when the blank count is zero, near zero (or there is a measurement involving a blank with zero variation). It is seen that for Poisson probability densities of the number of sample counts obtained, there is always a finite probability of obtaining a zero count when any small amount of activity is present. In fact, the probability of obtaining a zero count is exactly $\exp(-C)$, where C is the mean number of total counts in the counting time interval, as expected from the particular amount of activity present. Application of this formula shows that when $C = 3$, then $P(0) = \exp(-3)$ is 0.0497 (almost exactly 0.05). Currie recognized in his original paper that the constant 2.71 would not yield a Type II error of 0.05 for the very low background (or blank) situation. Thus, to keep the Type II error probability close to 0.05 as s_B approaches zero for low background measurements, as well as to simplify the L_D formula, the constant 3 is proposed for use in the MDA formulation in the bioassay performance standard. The formulation then becomes, for analytical procedures that are completed by a counting process:

$$\text{MDA} = \frac{4.65s}{KT}b + \frac{3}{KT}, \quad \text{Eq (8)}$$

where L_D has been converted to MDA (analogous to the conversion to LLD in Equation 1) by the use of a calibration constant, or conversion factor, that assumes that measurements have been converted to count-rate units in practice. Thus, K would be typically in units such as (counts per minute per microcurie), and s_b would be the standard deviation in the total number of counts of the blank in the total counting time interval. The standard deviation, s_b would be the square root of the total number of counts of the blank in the event that the only variability was the Poisson process variation; then, even though the square root is taken, the units of s_b would be counts, since $\mu = \sigma^2$ for the Poisson process.

For analytical processes that end in non-counting types of measurement of continuous variates, such as uranium fluorescence (where the number of light photons emitted per second is large enough to provide an almost continuous current signal), Eq (8) reduces to $4.65 s_b/K$, where s_b is the standard deviation, in a blank signal observed in the same manner as the analyte signal, and K is in appropriate units (such as $\mu\text{amp}/\mu\text{g.}$) Further cautionary notes and examples to aid in the use of Equation 8 will be given in following sections of this report. A more extensive discussion of the problem of determining MDA when s_b is not simply estimatable from Poisson statistics, and when there is considerable variation in K , has been presented recently by Currie (20).

It is important to realize in the application of Equation 8 that in almost any real batch of samples to be measured, there will be some distribution of activities present when in fact there are finite activities in the samples. Thus, the probability of a Type II error when L_C is taken as the decision level (actually L_C/KT in the same units as MDA), and an MDA quantity or more is present, is less than 0.05.

The formulation of Equation 8 has been validated for analytical methods using counting processes by observing that the Poisson process provides a 0.05 probability of a Type II error as s_b approaches zero, and that when the Poisson probability envelope has not sufficiently approached the Gaussian function for higher counts to allow the assumption of Normal probability densities over ranges in useful multiples of s_b , Equation 8 has been checked against the probabilities in Table 1 of Ginevan (21) to ensure that the formulation of MDA provides approximately a 0.05 probability of a Type II error, given that the decision level is taken to be:

$$L_C = 2.33 s_b \quad \text{Eq (9)}$$

Here, s_b has been substituted into Equation 4 as an estimate of σ_B for practical use in laboratory data analysis. In order to be a valid parameter for determining L_C and MDA, the value of s_b used in Equations 8 and 9 must either be known well from theoretical considerations and knowledge of measurement system stability, or s_b must be determined under current measurement conditions from a series of replicate measurements. Also, a paired blank count must be subtracted from each sample count to validate the conditions under which Equation 8 was

derived (7), unless for the purposes of quality control improvement, the laboratory wishes to subtract the value of a "well-known" blank (7) by taking very large counts of a suitable blank under stable conditions, which are assured to remain constant during the processes of sample analyses. Then, as indicated earlier in discussing Equations 8 and 9, the subtraction of a well-known blank for each sample analysis would tend to ensure 0.05 probability of a Type II error.

Although Ginevan's analysis (21) was carried out for use in comparing the health effects of two populations in an epidemiologic analysis, the case of counting samples and blanks when the total counts obtained are small is a close statistical analogue to the epidemiologic analysis, so his values in Table 1 of Reference 21 may also be used to estimate exact Type II probabilities for MDA formulations, or more exact MDA values than Equation 8 might provide when the Poisson Process is predominant. However, an examination of Ginevan's paper shows that, for purposes of this standard, the use of Equation 8 is sufficiently accurate for generic practical use in laboratory quality control performance testing. Further conditions on the determination of the factors in Equations 8 and 9 are discussed in the following sections.

2.4 Important Points and Examples in the Use of the MDA Formulation

2.4.1 Need to Use Appropriate Values of Each Variable

The NRC-DOE study at the Battelle Pacific Northwest Laboratories (22,23) of the appropriateness of the technical provisions of the draft ANSI N13.30 performance standard (24) showed in the early test runs that different laboratories had different methods of determining MDA's for analysis, even of the same nuclides. Some of these differences were necessary since there were different analytical methods or measuring systems employed in the analyses. In other cases, misapplication of MDA formulations, or use of insensitive methods, led to the inability of methods to meet the preliminary criteria for AMDA. These criteria have been adjusted in some cases in developing the standard (24) to be reasonable in terms of the radiation protection needs and considerations discussed in this report. In other cases, where the initial AMDA's were considered appropriate, it was apparent that guidance and standardized methods of determining MDA's were needed by the bioassay analytical laboratories. Many of the problems of determining MDA's (LLD's) have been examined recently by Currie (20) and have been discussed in detail in his report and bibliography. In this section of the report, some of the more important and frequently useful points and examples will be presented.

2.4.2 The Standard Error of the Count Rate

It may be useful to the user of Equation 8 to consider the ratio s_b/T as the standard error in the estimated blank count rate. This can be understood from the following considerations.

For a stable counting system, with purely Poisson variations, the standard deviation in a count C is given by:

$$s_b = C^{1/2}; \text{ and } \text{Var } C = C \quad \text{Eq (10)}$$

The count rate in counts per minute is

$$R = C/T \quad \text{Eq (11)}$$

From the rules of variance propagation,

$$\text{Var} (C/T) = \frac{1}{T^2} \text{Var} C = C/T^2 \quad \text{Eq (12)}$$

Therefore, taking square roots,

$$s_R = s_{C/T} = (\text{Var} (C/T))^{1/2} = C^{1/2}/T = s_b/T, \quad \text{Eq (13)}$$

which is the ratio found in the first term of Equation 8. Thus, the first term of Equation 8 may be considered as the ratio of the standard error of the count rate to the calibration constant, times 4.65; in other words, the first term is 4.65 times the standard error of the blank count rate in radioactivity units, or units in which the final analytical result is to be presented. This is the MDA when the count rate is Normally distributed.

2.4.3 Blank Signals with Gaussian Variations

When the final measurement of the analysis simulates a signal that is Gaussian (Normally distributed), such as the approximately steady current from a fluorimeter or the cumulative amount of charge produced by radon daughters in an ion chamber, the measured or calculated standard deviation in current or charge may be used in place of the standard error in count rate in Equation 8. Actually, the standard deviation in current or charge is analogous to a standard error in count rate, since any current or accumulation of charge is actually comprised of the effects of a large number of quantized electronic charges acting within the time interval of measurement observed in recording the analytical result. As pointed out by Currie (20), it is usually very difficult to calculate the standard error of the blank from basic principles of the analytical measurement; it is usually necessary to obtain an estimate of s_b (sometimes including the variations in K) from a limited number of experimental replications involving repeated measurements of a suitable blank. Variations in K, if determined separately, would have to be obtained by replicative measurements of samples containing the same composition as the suitable blank plus known quantities of the analyte under analysis.

2.4.4 Formulation of MDA for Variable Calibration or Conversion Factor K

If K varies significantly between measurements, or groups of measurements (for example, as a result of variations in chemical yield), the Working Group has concluded that it would be best to estimate the lower bound value that represents a 0.95 probability that values of K are higher than that bound, represented by $K_{0.05}$, and use $K_{0.05}$ in the denominator of Equation 8 when calculating MDA. This would ensure that the probability of a Type II error would be less than 0.05, even when fixed errors in recovery or calibration may affect an entire batch of samples. Then, at least less than 0.05 of the batches might be subject to the chance that Type II errors ("false negatives") could have a frequency greater than 0.05. This suggestion for dealing with variable K is similar to one proposed recently by Currie (see page 34, Reference 20) for handling variations in K.

In order to estimate the value of $K_{0.05}$, an evaluation of experimental and theoretical knowledge of the measurement system may be used to obtain the magnitudes of random and systematic errors, which may then be compounded with one of the suggested formulas such as (25):

"Uncertainty U is given at the 68 percent confidence level; that is,

$$U = (\sum_i \sigma_i^2 + (1/3) \sum_i \delta_i^2)^{1/2}, \quad \text{Eq (14)}$$

where δ_i are the estimated systematic uncertainties, and σ_i are the random uncertainties at the 68 percent confidence level..." (26).

The standard (25) also references two other methods (27, 28) of error analysis.

If an uncertainty U is calculated for K as in Equation 14, then $K_{0.05}$ could be taken as

$$K_{0.05} = \bar{K} - 1.645 U, \quad \text{Eq (15)}$$

where \bar{K} is the average value of K obtained from a series of calibrations, and the calculated MDA from Equation 8 would become

$$\text{MDA} = \frac{4.65 s_B}{K_{0.05} \bar{T}} + \frac{3}{K_{0.05} \bar{T}} \quad \text{Eq. (16)}$$

which would provide a conservative estimate of MDA taking into account possible systematic errors.

For example, if we assume that the sources of random error in K for use of the special Equation 1 are σ_E , σ_V , and σ_Y , these σ 's may be estimated from:

1. The repeated counting of a given laboratory standard source, whose activity is known by certification to within, say, $\delta_E = \pm 2\%$, to obtain an estimated standard deviation σ_E for the source-instrument system;
2. The estimation of random error, σ_V , in measuring sample volume V , which should be small; and
3. The estimation of σ_Y , when there are chemical yield variations or losses in obtaining the specimen to be counted for each bioassay sample, by repeated analyses of aliquots of a known spiked sample containing an amount of radionuclide sufficiently above the estimated MDA so that reasonably precise estimates of the yield Y for each aliquot can be obtained.

The values of δ_E , δ_V , and δ_Y can be estimated from knowledge of the maximum likely error in a standard source used repeatedly for counter calibration, the maximum error in volumetric determination from equipment specifications (this should be negligible for our purposes), and the maximum systematic error in yield that would probably be derived from the maximum error in the known activity concentration of the spiked stock solution used to determine chemical yields.

Finally, if there is an uncertainty in adding the interfering nuclide(s), or in preparing matrix materials that may affect self-absorption or background absorption of the "blank" specimen that is to be counted as the "appropriate blank," the coefficient of variation, CV, in preparing the overall blank should be estimated, and the value of U obtained from Equation 14 should be increased by CV·K, so that Equation 15 becomes

$$K_{0.05} = \bar{K} - 1.645 (U + CV \cdot \bar{K}) \quad \text{Eq (17)}$$

This correction would effectively account for contributions to s_b in Equation 8 that would not be apparent from simply replicating the counting process for a given blank specimen. Otherwise, if the analysis of individual sources of variation in the blank is impractical, the appropriate value of s_b in Equation 8 should be obtained by repeated counting of blank specimens that are prepared from the beginning of the blank-preparation procedure and placed individually into the final counting or measuring system for each blank specimen. However, as Currie points out, the method of replication would require about 200 replicates at each concentration to provide a relative uncertainty (95 percent confidence interval) in the s_b determination of $\pm 10\%$ (Reference 20, page 33).

Also, just to detect a systematic error equal in magnitude to the random error of a measurement process requires more than ten observations to reduce the standard error of the estimated measurement result (Reference 20, page 39). Therefore, it is best where possible to reduce non-Poisson variability in the measurement process to a minimum, particularly where it can not be estimated from theoretical or experimental knowledge.

It may be noted that, since K is in the denominator, errors in K that are due to biases of -33% to +33% will result in a bias error range of -25% to +50% both in the specification of MDA and in actual sample results. This latter bias range is the range currently proposed in the draft bioassay performance standard (24) for the average performance of finite test sample replicate in vitro measurements.

Currie (20) provides an excellent discussion and survey of the literature on additional considerations in estimating the MDA for various kinds of measurement processes.

2.4.5 Estimation of MDA in the Presence of Known Quantities of Interfering Nuclides

Pasternak and Harley (8) have shown, for the analysis of multi-component gamma spectra, how the LLD of a given radionuclide depends not only on the quantities of other radionuclides present, but also on the calibration spectra of all nuclides in the library. Their formulation of LLD for a given single nuclide without interfering nuclides that are not present in the "appropriate blank" is equivalent to that of Currie's L_D (7), insofar as the statistical model and probabilities of Type I and II errors. They have provided formulations of LLD for those cases where Poisson-distributed counts are large enough so that radionuclide activity estimators are approximately Normally distributed. The LLD's are considered in several categories in their paper:

1. The LLD for a nuclide when it is known that no other nuclides will be present and the library of calibration data contains only the spectrum for the radionuclide of interest;
2. The LLD for a nuclide when other nuclides are included in the library, as in a routine laboratory counting process, but other nuclides are not in the mixture; and
3. The LLD for a nuclide when other known nuclides are assumed to be present in the mixture at fixed levels, and the library contains calibration spectra for all nuclides involved.

Methods of Pasternak and Harley (8) may also be applied to alpha or beta spectra in principle, with some consideration for the differences in spectral shapes (9). If calibration spectra do not overlap, and the background or blank count is close to zero, then the $3/KT$ term of Equation 8 would determine the respective LLD for each nuclide.

More detailed guidance on methods of estimating LLD or MDA of a nuclide in the presence of the other nuclides is available in the literature, and reviewed in some detail in the more recent report by Currie (20). Since there is no single MDA in the presence of unpredicted quantities of interfering nuclides, the Working Group deemed that only the case of two nuclides would be useful for application in a standard where quality control testing of many laboratories, with different measurement systems, is specified. For the case of natural ^{40}K as an interfering nuclide, the range of variation in humans, or in human samples, is limited and rather well known. The average K-40 concentrations were recently measured (29) for 371 males to be 1.44 nCi/kg (± 0.404 nCi/kg) and for females to be 1.18 nCi/kg (± 0.410 nCi/kg). The K-40 concentrations versus age were (29):

$$\text{Males: K-40 (nCi/kg) = 1.67 - 0.00778 (Age) \quad \text{Eq (18)}$$

$$\text{Females: K-40 (nCi/kg) = 1.49 - 0.00977 (Age) \quad \text{Eq (19)}$$

$$\text{Both sexes combined: K-40 (nCi/kg) = 1.5 - 0.00895 (Age) \quad \text{Eq (20)}$$

These equations may be useful in determining some of the factors for Equation 17 to obtain $K_{0.05}$ for making a general statement of the MDA for a radionuclide where K-40 variability in the calibration matrix or in the blank, or phantom measurement corrections for the blank, are a consideration in the analysis.

In the case of whole body counting, when the shape of the K-40 spectrum is well-known, and when the ratio of the number of counts in the spectral region of interest to those in the K-40 photopeak for the phantom "blank" can be assumed to be well determined and applicable to the distribution of potassium in the subject, then the K-40 in the subject may essentially be used as its own control to obtain a paired blank. The standard deviation in the "blank" in this case would be:

$$s_b = R (C_{B-K-40})^{\frac{1}{2}} \quad \text{Eq (21)}$$

for the standard deviation of the blank count paired by the subject as his own control, where R is the ratio of the counts in the spectral region of the nuclide to be determined but caused by the (known) amount of K-40 in the phantom to the

number of counts in the K-40 photopeak for the phantom, for the total counting time interval of the subject, and $(C_{B-K-40})^{\frac{1}{2}}$ is the square root of the total counts in the subject's K-40 region.

When the shape of the K-40 spectrum in subjects is well predictable, and the photopeak of the nuclide to be measured is relatively narrow and lies on a slowly changing part of the K-40 spectrum, algorithms may be used that correct the measured photopeak counts for the nuclide to be determined by using adjacent parts of the spectrum to adjust for the amount of K-40 in the specific subject. This method would be most accurate for the case of a single nuclide whose photopeak is below the Compton edge of the K-40 spectrum, and would be difficult for a nuclide for which the photopeak is overlapping with the K-40 photopeak. If the algorithm involved, for example, a linear combination of three independent (statistically) terms proportional to the counts in three spectral regions, e.g., C_1 , C_2 and C_3 , and the correction to be subtracted from the analyte nuclide photopeak to obtain the result were of the form

$$f(C_1, C_2, C_3) = a_1 C_1 + a_2 C_2 + a_3 C_3 \quad \text{Eq (22)}$$

then the appropriate value of s_b could be estimated from the relationships

$$df = \sum \left(\frac{\partial f}{\partial C_i} \right) dC_i = a_1 (C_1)^{\frac{1}{2}} + a_2 (C_2)^{\frac{1}{2}} + a_3 (C_3)^{\frac{1}{2}} \quad \text{Eq (23)}$$

$$\text{Thus, } s_b = \sqrt{\text{Var}(df)} = \sqrt{a_1^2 C_1 + a_2^2 C_2 + a_3^2 C_3} \quad \text{Eq (24)}$$

However, this approximation would be applicable to Equation 8 only when each count C_i for the counting time interval is sufficiently large that Normal distribution statistics are approached, and so that the above chain differentiation involves small relative random errors in each C_i . If each C_i were greater than about 100 counts, then Eq. 24 could be used to obtain the appropriate blank value of s_b for Equation 8 for this type of data analysis.

Thus, it is clear, as pointed out in more detail by Currie (20), that the appropriate value of s_b for use in determining a detection limit may be determined not only by measurement error considerations, but also by considerations of the algorithms used in calculating the final result.

When both the spectral shape and the geometrical efficiency for detecting K-40 photons remains the same from sample to sample, as in a stable gamma spectrometry analysis of constant geometry urine sample residues, then the K-40 may be treated as an interfering nuclide using weighted least squares methods and algorithms as described by Pasternak and Harley (8). When there are instrumental instabilities, the computer programmed technique of Luizzi and Pasternak (30) or methods reviewed by Currie (20), may be used for transforming the calibration matrix to correct for baseline and gain discrepancies.

The weighted least squares solution for estimating nuclide concentrations $\theta' = (\theta_1, \theta_2, \dots, \theta_n)$ minimizes the sums of squares of differences between corresponding channel counts of the observed and "fitted" pulse height distributions,

when each squared difference is divided by an estimated variance of that difference calculated assuming Poisson statistics (the variance of each count equals the count). The minimizing of weighted squared differences, thus essentially minimizes the corresponding chi-square statistic of n degrees of freedom, assuming a "goodness of fit." Suppose that an n -channel background (not K-40 blank) corrected pulse-height distribution is obtained after counting a mixture of m ($< n$) distinct radionuclides. Let the distribution of the net counts in the n channels be represented by the row vector $x' = (x_1, x_2, \dots, x_n)$. The weighted least squares estimates of the nuclide concentrations present are given by the matrix equation (8):

$$\hat{\theta} = (AW^{-1}A')^{-1} AW^{-1} x \quad \text{Eq (25)}$$

where the $m \times n$ matrix A is called the "calibration matrix" and represents the set of standard or "library" spectra used in the analysis, and W is an $n \times n$ diagonal weighting matrix whose j^{th} diagonal element is

$$w_j = y_j/t^2 + b_j/r^2. \quad \text{Eq (26)}$$

In the above equations,

a_{ij} = the element in the i^{th} row and j^{th} column of the matrix A ;
viz. the average net-count rate in channel j per unit amount
of nuclide i , obtained from a "large" number of measurements;

y_1, \dots, y_n = gross counts observed in channels 1..... n ;

b_1, \dots, b_n = background count observed in channels 1 through n ;

x_1, \dots, x_n = net-count rate observed in channels 1 through n ;

t = sample counting time; and

r = background counting time.

Equation 26 essentially is an estimate of the j^{th} "weighting factor," which is the Poisson variance in the net count rate x_j in the j^{th} channel. One notes that this analysis allows for a different sample counting time, t , and background counting time, r .

The variance of the estimated concentration parameter for the i^{th} nuclide may be taken (8) as the i^{th} diagonal term of the matrix $(AW^{-1}A')^{-1}$.

One of the examples given by Pasternak and Harley (8) of the use of the weighted least squares method is the two-channel case when two nuclides are present. This case can be applied to the whole body counting analysis for a single nuclide (or at least a nuclide not interfered with by nuclides having photopeaks in lower ranges) in the presence of varying and unknown amounts of K-40. The spectral region can be divided into two ranges of channels to simulate the following two-channel, two-nuclide case. In this case, the standard errors of the estimated concentrations of nuclide 1 and nuclide 2 (e.g., K-40) are shown (8) to be:

$$SE(\hat{\theta}) = \left\{ \frac{1}{d} \left(\frac{a_{21}^2}{y_1/t^2 + b_1/r^2} + \frac{a_{22}^2}{y_2/t^2 + b_2/r^2} \right) \right\}^{-\frac{1}{2}} \quad \text{Eq (27)}$$

and

$$SE(\hat{\theta}) = \left\{ \frac{1}{d} \left(\frac{a_{11}^2}{y_1/t^2 + b_1/r^2} + \frac{a_{12}^2}{y_2/t^2 + b_2/r^2} \right) \right\}^{-\frac{1}{2}} \quad \text{Eq (28)}$$

where d denotes the determinant of the matrix $AW^{-1}A'$.

As the shape of the two-channel distribution of the second nuclide approaches that of the first, $d \rightarrow 0$ and the standard errors of the estimates of the nuclide concentrations approach infinity. Thus, the estimation of individual nuclide concentrations for two nuclides whose spectra completely overlap is found to be impossible, as may be expected (8).

The LLD for nuclide 1 (and MDA in the terminology of this standard) is approximated by (8).

$$LLD_1 \cong (k_\alpha + k_\beta) SE(\hat{\theta}_1) \quad \text{Eq (29)}$$

where $SE(\hat{\theta}_1)$ is obtained from Equation 27, and the LLD for nuclide 2 is approximated by

$$LLD_2 \cong (k_\alpha + k_\beta) SE(\hat{\theta}_2) \quad \text{Eq (30)}$$

where $SE(\hat{\theta}_2)$ is obtained from Equation 28. As in Currie (7), k_α and k_β are the multipliers that set the LLD at an amount where the probability of a Type I error (false positive) is 0.05 and the probability of a Type II error (false negative) is 0.05, when the decision that something is present is made whenever the quantity $\hat{\theta}_i$ obtained from the count by the algorithm is greater than or equal to $k_\alpha SE(\hat{\theta}_i)$.

Since $SE(\hat{\theta}_1)$ for the estimated quantity of nuclide 1 in Equation 27 depends on the gross counts y_1 and y_2 in both channels (or spectral ranges) and on the counts a_{21} and a_{22} contributed by the nuclide 2 standard source to channels 1 and 2, respectively, the LLD for nuclide 1 obviously depends on the quantity of nuclide 2 present and on the sensitivity of the detection system to gamma rays from nuclide 2. Similarly, the LLD of nuclide 2 is dependent on the quantity of nuclide 1 present and its calibration spectrum. Thus, an LLD (or MDA) for a nuclide measured in the presence of K-40 must be specified as applicable to a given quantity of K-40 in a given phantom-geometry, or sample-geometry, arrangement.

For a given quantity and geometry of K-40 as nuclide 2, a LLD (or MDA) for another nuclide 1 may be estimated by providing a simulated unknown sample having zero activity of nuclide 1 but the desired activity of K-40, determining for Equation 27 the background counts b_1 and b_2 for the ordinary background counting times r , determining y_1 and y_2 for the ordinary sample counting times t_1 and

t_2 , determining a_{21} and a_{22} (the net count-rates in channels 1 and 2, respectively) accurately from long counting times of the "standard" known quantity of K-40 in a simulated unknown with zero activity of nuclide 1 (and no other interfering activity), and then using Equations 27 and 29 to determine the LLD (or MDA).

Rogers (31) discusses the development of an algorithm to calculate the LLD's for ^{90}Sr and ^{89}Sr when they are measured simultaneously (31, 32). Currie (Reference 20, pages 108-111) presents the formulations of detection limit concepts for multi-channel weighted-least-squares analysis of two-nuclide spectra, including suggestions for accounting for systematic errors in baseline shape and level. Expansion of Equation 25 and its variance-covariance matrix for more than two nuclides and two channels would require a lengthy presentation and is beyond the scope of this report. The corresponding calculations would also be extremely tedious, and thus usually require appropriate computer programs (8, 20, 30).

Currie (Reference 20, pages 135-136) suggests using the output from a weighted least squares spectrum convolution to estimate an upper limit estimate of the LLD of a nuclide in the presence of the amounts of other nuclides found. If the estimated concentration and standard error of nuclide 1 were $\theta \pm \text{SE}_1 = 95.6 \pm 32.2$ pCi/L, as obtained from a specific computer program, the standard error for zero activity of nuclide 1 (in the presence of the specific quantities of other nuclides) would be something less than 32.2 (since Poisson variance decreases with the mean). Therefore, Currie points out, neglecting systematic error, the upper limit of the decision level concentration and the upper limit of LLD would be (20):

$$X_c' = 1.645 \sigma_x = 53.0 \text{ pCi/L} \geq X_c \quad \text{Eq (31)}$$

and $X_D' = 2X_c' = 106 \text{ pCi/L} > X_D$

Thus, the above example result of 95.6 pCi/L would be judged to be significant (detected) and 106 pCi/L could be taken as a (conservative) upper limit LLD in the presence of the particular observed quantities of other nuclides.

2.4.6 Currie's MDA (LLD) Formulation, Including Adjustments for Limited Systematic Errors in Blank and Calibration Factor

One of the general formulations for detection limit suggested by Currie takes into account blank and calibration systematic errors (Reference 20, pages 38-40, pages 78-79, and pages 92-94). This formulation will be summarized here for its potential usefulness in dealing with such systematic errors.

If Δ represents a bound on the absolute systematic error of determining a blank (or interference) count, and f is a proportional correction for systematic error in the calibration factor K ; then the decision level and detection limit may be written as (20):

$$S_c = \Delta + Z_{1-\alpha} \sigma_0 \quad \text{Eq (33)}$$

$$X_D = f(2\Delta + Z_{1-\alpha} \sigma_0 + Z_{1-\beta} \sigma_D) / KT, \quad \text{Eq (34)}$$

where f may be taken as the relative systematic error in K :

$$f = 1 + \Delta_K \quad \text{Eq (35)}$$

and the relative systematic as well as random error in K may be estimated from the random and systematic errors in such factors as E , V , and Y (defined for Equation 1) by a formula such as Equation 14. If the systematic errors are small, Δ_K may be taken as the fractional random error (coefficient of variation), ϕ_K , where

$$\phi_K = (\phi_E^2 + \phi_V^2 + \phi_Y^2)^{\frac{1}{2}}, \quad \text{Eq (36)}$$

where all values of ϕ are relatively small ($\ll 1$).

If we set $k_\alpha = k_\beta = 1.645$ as in most suggested MDA formulations, let Δ_B equal the estimated maximum fractional systematic error in the blank count B (with the blank containing the specific quantities of "interfering nuclides" when present), and let $K = 2.22$ YE \bar{V} as in Equation 1 for converting counts of a radiochemically separated sample into pCi/L, then Equations 33 and 44 become:

Decision Level:

$$S_c = \Delta_B B + 1.645 \sigma_o \quad \text{Eq (37)}$$

and

Detection Limit (MDA or LLD):

$$x_D = (1 + \Delta_K) (2\Delta_B B + 1.645 \sigma_o + 1.645 \sigma_D) / KT \quad \text{Eq (38)}$$

for $B \geq 70$ counts (19).

where σ_o is the standard deviation in the net count of a sample that has no activity of the analyte, and σ_D is the standard deviation of the net count of

a sample containing the LLD quantity of the analyte as well as the blank composition.*

This is what Currie (Reference 20, page 34) calls the "S-based" approach, where the decision of whether or not something is detected is made purely in the signal domain on the basis of a certain deviation in a sample count or measurement from an expected blank count. The MDA (or LLD) that is defined to provide enough counts so that a Type II error (false negative decision) is made only 5% of the time (or less) will depend also on the calibration factor K, which may have appreciable systematic or random errors, or both. These errors need to be accounted for by the term Δ_K when they are relatively small (say, $\theta_K/K < 0.1$), or by methods such as those discussed in section 2.4.4.

If K is taken as 2.22 EVY, with $T = t_1$ the sample counting time in minutes, and other factors defined as in Equation 1, then a numerator in units of counts in Equation 38 can be converted to an MDA (or LLD) in units of pCi/L, for a measured sample separated from an initial matrix of volume V. If the sample of interest is simply a planchet of material, the volume V may be omitted to obtain the MDA (or LLD) in terms of activity units on the planchets that would have only a 5% chance of nondetection.

*The values of σ_0 and σ_D would in practice need to be replaced by some kind of sample-measured standard errors s_0 and s_D as estimates. The s_0 value would be obtained from

$$(s_{B_1}^2 + (t_1/t_2)^2 s_{B_0}^2)^{1/2} = (B_1 + (t_1/t_2)B_0)^{1/2}$$

when the errors are essentially the Poisson errors in total counts; here, the standard error s_0 is obtained as it would ordinarily be calculated from a routinely measured sample (that happened to be a "blank" because no additional analyte beyond that in the "appropriate" blank was present), counted for time t_1 , with the resulting count B_1 being corrected by subtracting a count $(t_1/t_2) B_0$ obtained from B_0 counts in time t_2 for a paired "standard" blank, routinely used for the analysis. The value of the standard error of the net sample count when an LLD quantity is counted would be s_D , used to estimate σ and similarly obtained as $(C_D + (t_1/t_2)B_0)^{1/2}$, where C_D is the gross count in time t_1 for a sample containing an LLD quantity of the analyte nuclide added to a matrix containing the composition of the blank. When random errors additional to those contributed by natural Poisson counting fluctuations are present, they should be estimated by counting replicate blanks and LLD quantities (after LLD has been estimated to a first approximation), and the values of s_0 and s_D should be obtained from

$$(s_{B_1}^2 + (t_1/t_2)^2 s_{B_0}^2)^{1/2} \text{ and } (s_{C_D}^2 + (t_1/t_2)^2 s_{B_0}^2)^{1/2}$$

respectively.

If the sample of interest is a person to be counted for total body Cs-137 in a whole body counter, then the constant K might be the net count obtained in $T = t_1$ minutes in the Cs-137 photopeak region per unit activity of Cs-137 distributed with an average K-40 content per unit activity, in a suitable phantom, after subtracting the expected counts in the Cs-137 photopeak spectral region from Compton-scattered K-40 photons. The Compton-scattered contribution from K-40 may be estimated from a separate measurement of the K-40 scatter contribution using a phantom containing the same amount of K-40, but no Cs-137. Estimated maximum systematic uncertainties in K, due perhaps to such biases as self-absorption and geometry differences resulting from a person's torso being a different size and shape and sagging into a reclining position during measurement (which may produce random errors from person to person as well as systematic biases), may be incorporated in the Δ_K term of Equation 38. Uncertainties due to baseline or amplifier gain shifts in the multi-channel analyzer could be incorporated into the Δ_B terms of Equations 37 and 38.

To obtain estimates of σ_0 for whole body counting of a person in the above example, the K-40 scatter contribution from the phantom may be corrected to take into account age, height, weight and sex according to methods reported in the literature. Alternatively, the K-40 scatter contribution in the Cs-137 photopeak region may be estimated by extrapolation or fitting of the nearby K-40 spectrum of the person being measured in regions where other nuclide contributions do not interfere, as described in section 2.4.5 or in Currie (20). With an estimated count B in time T for the person's K-40 scatter contribution in the Cs-137 photopeak region, the estimate of σ_0 may be taken as:

$$\hat{\sigma}_0 = (\hat{B} + C^2 K_{40})^{1/2}, \quad \text{Eq (39)}$$

where C = the correction applied to the counts K_{40} in the potassium-40 photopeak part of the phantom spectrum to obtain the "blank" counts due to K-40 photon scatter contribution in the Cs-137 photopeak region of the patients, and B is the estimated blank count,

$$\hat{B} = C K_{40} \quad \text{Eq (40)}$$

(The C^2 term occurs in Equation 39 because $\text{Var}(C K_{40}) = C^2 \text{Var} K_{40} = C^2 K_{40}$.) Equation 39 takes into account the fact that, for a person with no Cs-137, the (variable) count approximating B in the Cs-137 photopeak region would be compared with (by subtraction) the (also variable) estimate of B obtained from $C K_{40}$. Ideally, the errors in C would be relatively negligible due to prior careful and precise K-40 spectral shape determinations.

In general, $\sigma_D \neq \sigma_0$, for an LLD quantity of Cs-137 present, since the additional counts in the Cs-137 photopeak region also add variance to the variance in B. An accurate determination of σ_D for an LLD quantity of Cs-137 present would thus, of course, require a priori knowledge of the LLD, which generally does not exist. Thus, an approximation of $\sigma_D = \sigma_0$ can be used in Equation 38 to obtain a first approximation to the LLD. Then, this quantity could be added to a suitable phantom, repeated counts taken until B is at least 70 each time (20), and then the standard deviation in the net count in the Cs-137 photopeak region taken as σ_D :

$$\hat{\sigma}_D = \left(\sum_{i=1}^N \frac{(C_{Di} - \bar{C}_D)^2}{N-1} \right)^{1/2}, \quad \text{Eq (41)}$$

since the correction CK_{40} cancels in each difference term in the numerators (unless paired K_{40} measurements are made for each C_{Di} measurement).

Using estimated values of $\hat{\sigma}_0$ and $\hat{\sigma}_D$ from Equations such as 39 and 41 in Equation 38, as well as estimated systematic error corrections Δ_K and Δ_B , will then give a more reliable estimate of MDA (or LLD) in the sense that confidence will be provided that the chance of a Type II (false negative) error will probably be less than 5 percent.

Simplified LLD formulation for Blank Counts ≥ 70

For a large enough count B (e.g., ≥ 70) so that the constant term 2.71 can be ignored and $\sigma_0 \sim \sigma_D$, and systematic errors that can be assumed bounded by $\Delta_B = 0.05$ and $\Delta_K = 0.1$ (5% baseline error and 10% calibration factor error),

Currie (20) has derived a simplified LLD formula somewhat as follows:

$$x_D = 1.1(2(0.05)B + 3.29\sigma_0)/KT, \quad \text{Eq (42)}$$

which becomes

$$x_D = \frac{0.11B}{KT} + \frac{3.62\sigma_0}{KT} \quad \text{Eq (43)}$$

If a (near-blank) sample is measured for time t_1 yielding y_1 counts, and a blank is measured for time t_2 to get y_2 counts, the net count (signal S) of a near-blank sample would be (Reference 20, page 92):

$$S = y_1 - \left(\frac{t_1}{t_2}\right) y_2 \quad \text{Eq (44)}$$

The variance in S is:

$$\text{Var } S = \text{Var } y_1 + \text{Var} \left(\frac{t_1}{t_2} y_2\right) \quad \text{Eq (45)}$$

$$= \text{Var } y_1 + \left(\frac{t_1}{t_2}\right)^2 \text{Var } y_2 \quad \text{Eq (46)}$$

For $\text{Var } y_1 = B$, and $\text{Var } y_2 = \text{Var} \frac{t_2}{t_1} B$, since here $t_2 B/t_1$ is

a directly-measured poisson variate, so that

$$\hat{\sigma}_0 = \sqrt{\text{Var } S} = \sqrt{\hat{B}(t_1 + t_2)/t_2} \quad \text{Eq (46)}$$

Substituting from Equation 46 into Equation 43 yields:

$$x_D = \frac{0.11B + 3.62 \sqrt{\hat{B}(t_1 + t_2)/t_2}}{KT} \quad \text{Eq (47)}$$

which is equivalent to Equation 24, page 93, Reference 20*. Here K may be 2.22 EVY for nuclide concentration analyses or a suitable calibration factor for other measurements. The counting time of the sample is $T = t_1$, and the (paired) background is assumed to be counted for time t_2 . B is the blank count in time $T = t_1$.

The critical number of counts at which a decision is made that there is activity present remains as simply (see Reference 20, page 92):

$$S_c = 1.645 \sigma_0 = 1.645 \sqrt{\hat{B}(t_1 + t_2)/t_2} \quad \text{Eq (48)}$$

2.4.7 MDA Formulation for Rapidly Decaying Nuclides

As Currie has pointed out (Reference 20, p. 75, p. 79, p. 105, p. 137), when the radioactive decay of a nuclide appreciably decreases the quantity of radioactivity during the time period(s) of measurement, additional corrections are needed in the denominator of the LLD equation. For a single counting time, the value of T in the denominators of Equations 8 and 47 may be replaced by:

$$T_c = e^{-\lambda t_a} (1 - e^{-\lambda \Delta t}) / \lambda, \quad \text{Eq (49)}$$

where λ = the radiological decay constant of the analyte nuclide (in units of min^{-1} , if the 2.22 factor is used in K to convert to pCi/L)

t_a = the time interval (usually in minutes) between the time of interest for calculating the quantity of nuclide present and the initiation of the measurement (counting) process,

Δt = the duration of the measurement (counting) interval.

Of course, the simple Equation 49 would not be applicable if interfering nuclides were also decaying, or if the measurement process involved multiple separations and counting intervals. No single expression can be given for multiple counting schemes, since they depend upon the exact design (and input function) for the sampling or counting process (Reference 20, p. 75).

Currie also points out that for the measurement of a nuclide of finite half-life, both the half-life and the systematic error bounds limit the amount of LLD

reduction that can be accomplished by increasing the counting time. For a "typical" ^{131}I example with 1% assumed interference systematic error and 10% assumed systematic calibration error, Currie shows that increasing counting time from 200 minutes (giving a $B = 6000$) to 2 weeks, a 100-fold increase in counting time decreases the overall LLD only 25 percent (Reference 20, p. 137).

2.4.8 MDA Determinations When the Physical Processes of Measurement are Stable and Well Known

As Currie pointed out, there is no all encompassing rigorous solution to the problem of non-Poisson random and systematic error effects on detection capabilities (Reference 20, page 74). Also, Currie notes that carefully controlled experimental work is necessary to even approximate a situation where Poisson errors of known distribution allow LLD determinations that are meaningful in terms of Type I and Type II errors that are close to the intended probability levels (Reference 20, page 38). Nevertheless, measurement systems may be available that are stable and have signals whose distributions are predictable from basic physical theory.

For example, consider small samples of Am-241 measured in a deep and thick NaI well counter, whose quenching and measurement system stability is regularly monitored with an internal gamma-ray source. The ultimate pulse height distribution in the 60-keV photopeak region may be predictable from the sequence of known physical events all the way from the Poisson nature of the photon emission, to the known probabilities of interaction and electron-energy distributions of Compton and photoelectric events, which produce light pulses and photomultiplier pulses whose pulse-height distributions can be characterized.

The mean and variance of the ultimate pulse-height distribution of such a system, and even the mean and variance of a current signal from the photomultiplier in a specified time interval (33), may be calculated theoretically by adapting the generating function approach of Seitz and Mueller (34) for determining the means and variances of the resulting signals of the measurement process. These means and variances can then be used to calculate s_b , s_D and the respective MDA's from Equation 8 or 47.

For most laboratories and measurement systems, however, it would probably be more practical to determine the values of s_b and s_D from replicate measurements of blanks, and blanks-plus-relevant quantities of the analyte, respectively (20). When systematic errors are significant, it may be necessary to determine the possible influence of both systematic and random errors at various concentrations from none added above the blank to several times MDA (20).

2.4.9 MDA Formulations for Measurement Processes Having Continuous Output Signals

As indicated in Sections 2.4.2 and 2.4.3, a continuous signal detector will have a standard error in its current, or signal, as observed for the given observation time, that may replace s_D/T in Equation 8. The longer the observation time, even by the human eye, the lower the amount of difference in current that can be distinguished from a given background or blank signal, with its concomitant variance.

If the net standard error σ_D in observing the continuous signal (during a standard observation time or observation process) for an LLD (plus blank) quantity present, differs from the net standard error σ_o of measuring a sample that is effectively a blank relative to the blank alone, then the LLD (or MDA) may be obtained from the equation:

$$\text{MDA} = f \frac{(\Delta + 1.645 \sigma_o + 1.645 \sigma_D)}{K}, \quad \text{Eq (50)}$$

where

K = the calibration constant for the measurement in units of current or signal per unit of activity or mass (e.g., microamps signal averaged over an observation time Δt per μg uranium);

Δ = the absolute maximum systematic error in the signal blank or interference determination in signal units

f = the estimated maximum relative error in the calibration constant K , plus 1, (i.e., $f \geq 1$).

Here, for use in Equation 50,

$$\hat{\sigma}_o = \sqrt{s_b^2 + s_{BE}^2} \quad \text{Eq (51)}$$

where s_b = the calculated standard error in measuring blanks, and

s_{BE} = the calculated standard error in measuring blank-equivalent samples, where no added analyte is present in the blank medium but the measurement process has a different $s_{BE} \neq s_b$ for the blank equivalent sample.

Ordinarily, $s_b = s_{BE}$, so that Equation (51) would give:

$$\hat{\sigma}_o = \sqrt{2} s_b \quad \text{Eq (52)}$$

Similarly,

$$\hat{\sigma}_D = \sqrt{s_b^2 + s_{b+D}^2}, \quad \text{Eq (53)}$$

where s_b = the calculated standard error in measuring blanks, and

s_{b+D} = the calculated standard error in repeated determinations of the continuous signal with an LLD + blank quantity present.

In the special case where the standard error of the blank is approximately equal to the standard error s_{b+D} with a blank-plus-an (initially approximated) LLD quantity present, then

$$\hat{\sigma}_D \cong \sqrt{s_b^2 + s_b^2} = \sqrt{2} s_b \quad \text{Eq (54)}$$

If both Equations 52 and 54 are valid, $f \cong 1$, and $\Delta = 0$, then Equation 50 may be simplified to become:

$$\text{MDA} = \frac{4.65 s_b}{K}, \quad \text{Eq (55)}$$

for the case where blank variability affects the variance of the calculated net sample activity.

If, on the other hand, the signal of a sample measurement (even a sample that was "blank-equivalent", having no added quantity of analyte) had an $s_D = s_b$ for a blank measured for a given observation time, and was compared with a "well-known" blank where $s_b = 0$, then

$$\begin{aligned} \text{MDA} &= \left(\frac{1.645 \sigma_o + 1.645 \sigma_D}{K} \right) \\ &= (1.645 \sqrt{0 + s_b^2} + 1.645 s_b) / K \\ &= \frac{3.29 s_b}{K} \quad (\text{for a well-known blank}) \quad \text{Eq (56)} \end{aligned}$$

It would be appropriate to use Equation 56 in a situation where the pointer on a meter with a sample in place, with a fluctuation characterized by s_D , is being compared in its average position with a well-known line on the meter, whose measurement of a standard well-known blank can be assumed to be absolutely stable. In the more common situation, the position of a fluctuating pointer on a (continuous) meter scale is being judged for its difference in position compared to the fluctuating position of the pointer when a paired blank (or background) was in position in the measuring instrument. In the latter case, detection of differences would be more difficult so it would be more appropriate to calculate the MDA from Equation 55 (or the more general Equation 50 if s_D can be characterized and is different from s_b , and if f and Δ can be specified).

The considerations in the above paragraph are applicable whether the observation of the pointer is by human eye, estimating an average pointer position over some given time interval Δt , or whether the observation of the pointer position is by some more objective measurement process such as one that integrates charges over a time interval Δt and provides a digital numerical display. Evidently, careful consideration, and perhaps some investigative experiments and measurements, must be applied before a decision can be made on the applicable formulation of MDA for a continuous (as well as for a discrete) measurement process.

When the standard errors in continuous rate-meter measurements may be attributed primarily to Poisson fluctuations in count rates of stable blanks or sample radionuclide quantities, standard errors may be estimated by the formulation summarized by Currie (Reference 20, pages 51 and 96). If the observation time t is long compared to the system time constant T for an analog circuit, the standard error in count rate may be obtained from:

$$\sigma_R^2 = R/t, \text{ if } t \gg \tau \quad \text{Eq (57)}$$

On the other hand, if the time constant is long compared to the observation time t (an effectively "instantaneous" measurement), then

$$\sigma_R^2 = R/2, \text{ if } t \ll \tau \quad \text{Eq (58)}$$

However, the use of Equation 58 assumes that the instantaneous observation of the count rate is made after the pointer reaches a stable position after a time long compared to τ , during which the instrument is exposed to a constant blank or amount of sample radioactivity. Equations 57 and 58 may be used to obtain standard errors in count rate for either the sample or the blank, when Poisson fluctuations predominate.

2.4.10 Measurements in Uncontrolled Environments

As Currie has pointed out (Reference 20, page 96), an instrument such as a count-rate meter may be subject to rather significant non-Poisson "background" variations in an "uncontrolled" environment. For example, if an incompletely shielded detector in a counting laboratory is subjected to background variations contributed by changes in the air concentrations of natural radon or thoron over time (due perhaps to incomplete filtration in the ventilation system), and this is substantiated by some goodness of fit test on a sequence of background counts, then the maximum expected standard deviation s in background or blank count rate should be determined and used in Equation 55 for the determination of MDA, with a corresponding equation such as

$$L_C = 2.33 s \quad \text{Eq (59)}$$

for the decision level signal.

2.4.11 Multiple Detection Systems for More Than One Nuclide

Another cautionary note in regard to the fact that systems that search for and measure the presence of more than one nuclide at a time may have probabilities >0.05 of a Type I error due to multiple sequential application of decision levels such as Equation 4 or 59. This phenomenon, not often recognized, requires appropriate adjustment of the decision level for such systems before a decision is made that radioactivity is present above that in the blank (Reference 20, pages 14, 64-66, 113). Care in stating the results of such multiple detection measurements is necessary to avoid serious misinterpretations of multi-component spectral data.

2.4.12 Detection By Visual Observation

In a 1977 IAEA test intercomparison of the ability of 200 participating laboratories to detect, resolve and evaluate high resolution Ge(Li) gamma spectra by gamma-ray peak evaluation algorithms, while most participants were able to produce results with the easily detectable single peaks, less than half of them provided reliable uncertainty estimates (Reference 20, pages 60-62). Two-thirds of the participants attacked a problem of resolving nine doublets, but only 23 percent were able to provide a result for the most difficult case. Accuracy assessment for the doublets was unreliable. For a peak detection exercise, with

22 subliminal peaks, the number correctly detected ranged from 2-19. Large numbers of false positives occurred, ranging up to 23. Considering the modelling and computational power available, it was interesting that the best peak detection performance was given by the "trained eye" (Reference 20, page 64).

Thus, it is important for quality control and observation of any anomalies of measurement that spectra be visually examined, and original data of measurement be examined, by the analyst. Often the human eye of a knowledgeable observer can detect nuclides, or the suspicion of nuclides, that are not detectable using the algorithms designed for systematic analysis.

2.4.13 Examples of MDA Calculations

In this subsection, some examples of typical and simple MDA calculations will be given.

Example for Uranium Fluorimetric Measurement

Standard deviation of repeated measurements of 0.2 ml blanks for unexposed individuals = ± 0.061 microamps (for a mean of 0.0725 μA). $K = 287 \mu\text{A}/\mu\text{g}$ $\text{U}_3\text{O}_8/0.85 = 337 \mu\text{A}/\mu\text{g}$, from spikes of 0.012 μg U_3O_8

$$\text{MDA} = 4.65 \frac{0.061}{337 \times 0.0002\text{L}} = \underline{4.2 \mu\text{g U/L}}$$

E is assumed = 1 for sample relative to blank fluorescence measurement in fusion dish, and chemical yield is assumed = 1 since the entire 0.2 ml of aliquot is placed into the fusion dish.

Example for Plutonium Alpha Counting

1. Background = 1200 alpha counts/200 min = 6 c/min
Efficiency, E, of counting = 0.46; Average chemical recovery = 0.60

$$\text{MDA} = 4.65 (1200)^{\frac{1}{2}}/0.60 \times 0.46 \times 200 = \underline{6 \text{ dis/min}},$$

(for procedures involving 200 minute counts of blank and of sample)

NOTE: We assume here that the term $3/0.60 \times 0.46 \times 200 = 0.054 \text{ d/min}$ is negligible.

2. Background = 0 alpha counts/200 min, other parameters as above

$$\text{MDA} = 3/0.60 \times 46 \times 200 = \underline{0.05 \text{ dis/min}}$$

NOTE: MDA now is inversely proportional to T.

Example of Whole Body Counting of Am-241

Blank count of phantom with normal K-40 at one meter below 5" x 4" NaI crystal in standard chair = 400 counts in 10 minutes under 60 KeV area.

Relative standard deviation due to variations in placement of detector and phantom between blank and calibration measurements = 25%.

Calibration count with 2 microcuries of Am-241 (plus K-40) in phantom
= 4,000 counts/minute in 10 minutes (net count after blank subtraction)

MDA = 4.65 S/KT, where $S = 1.25 s_b$ to correct for placement error

$$\begin{aligned} \text{MDA} &= 4.65 (1.25 (400)^{\frac{1}{2}}) / (4,000 \text{ c}/10\text{min}-2\mu\text{Ci}) 10 \text{ min} \\ &= \underline{0.058 \mu\text{Ci}} \text{ (at one meter from standard chair)} \end{aligned}$$

(Assumes K-40 scattered photons and other interfering activity are the same in the subject and in blank; an appropriate correction for K-40 activity differences in the phantom and in a subject who is measured should be made to determine the MDA applicable to the particular subject (see Section 2.4.5).

Uranium-235 Lung Counting

Lung counting for the gamma lines of U-235 was carried out for 22 employees of NLO, Inc., beginning in 1968 and carried out for many years following retirement (35). Since for each employee visit to the in vivo lung counter, triplicate measurements were made, the data provided an excellent basis for expressing relative errors of measurement versus lung burden, and offer an example of how in vivo measurement error can vary with quantity of nuclide to be determined. The coefficient of variation of these determinations varied inversely as the quantity of U-235 present, with the 95% confidence interval of each measurement approximating $\pm 30 \mu\text{g}$ U-235, and thus $\hat{\sigma}_D$ and $\hat{\sigma}_b = \pm 15 \mu\text{g}$, in absolute value, both relatively independent of amount of U-235 present. Since the specific activity of U-235 is $2.14 \mu\text{Ci/g}$, and the $s_b = \pm 15 \mu\text{g}$ here includes all errors of determination as estimated from many replications, the MDA may be estimated directly as

$$\begin{aligned} \text{MDA} &= 4.65 (\pm 15\mu\text{g} \times 2.14 \mu\text{Ci/g} \times 10^{-6}) \\ &= 150 \times 10^{-6} \mu\text{Ci} = 0.150\text{nCi}, \end{aligned}$$

which is in approximate agreement with the value of 0.2 nCi initially selected by the judgment of the ANSI N13.30 Working Group (24).

Plutonium Lung Counting

The determinations of MDA for lung counting to quantities of plutonium and some transplutonium nuclides in the human lung involve many special problems, such as the determination of the influence of chest wall thickness on the absorption of very low-energy X-rays (12-25 KeV region), the estimation of the added contributions to background from "escape peaks" caused by X-rays excited by the analyte that escape the detector, and differences in distribution of the analyte within the subject's lung from that in the phantom (blank) lung (36-39). Examples of analyses or corrections for these sources of uncertainty may be found in the literature for many types of measurement systems. Currently, MDA's for these nuclides are not very much below the quantities that may be considered to deliver maximum permissible doses if present in the body for extended periods

of time (40, 41). The best MDA's for plutonium in the lungs for 2000 second counts are about 4 nCi for ^{238}Pu , and about 10 nCi for ^{239}Pu since the L X-ray intensity of ^{239}Pu is less than half that of ^{238}Pu (36). Thus, careful determination of the MDA's for such measurements usually requires a detailed experimental investigation of all physical and biological factors that may influence detection and measurement.

2.5 Recording of Analytical Results

The development and formulation of MDA concepts, for the purpose of properly representing, a priori, measurement system capabilities (the assurance of high probabilities of detecting given amounts) should not be misconstrued and misused for purposes of rounding off - and thus biasing - a posteriori measurement results. A result below the stated MDA does not, all-of-a-sudden, lose all information content, just as a result above MDA is not perfectly precise. If a quantity L_C of about $\frac{1}{2}$ MDA were actually present (20), it would have a 50% probability of being called "non-detected" or "insignificant" - yet it is actually there; so the probability of a Type II error would be 50% if the measurement were normally distributed.

Thus, information is present below MDA, so results should be recorded as the best estimate with specifically stated confidence intervals about the estimated mean, even if the mean has a negative value (Reference 20, pages 7, 24, 32, 58, 95). If a "less-than" value is desired at a given confidence level, it should be separately estimated for each result, not taken as L_C/KT or MDA, and not generally used in place of the recorded result and its confidence interval (7, 20).

Analytical results and their confidence intervals may be estimated by many methods or algorithms available for either single or multi-component analyses (8-10, 20, 21, 30-32, 39, 42).

The recording of external doses below a minimum detectable amount of 30 mrem as exactly 30 mrem has been estimated to bias an entire segment of population exposure by more than a factor of 3 too high (43). Also, the minimum detectable amounts of measured external doses have sometimes been overestimated by factors of 3 or more, so that long-term average doses, or cumulative population collective doses may be very seriously underestimated or overestimated by improper recording of data (44, 45). Even though internal dose estimates from bioassay data may have overall uncertainties of factors of 3 or more due to inherent random as well as systematic errors (46), the addition of bias by improper data recording could result in serious errors in long-term assessment of individual and population exposure (see Section 3). Sometimes when data are initially recorded in a proper manner to preserve information, later knowledge gained about the measurement process, or biological models used to estimate internal dose, can be used to retrospectively correct earlier data sets to minimize the bias in the long-term (and often more important) dose assessments. In addition, random errors often tend to cancel with an increasing trend toward better precision over the long term, in accordance with statistical laws of large numbers.

2.6 Considerations in Selecting Representative MDA's

The specific reasons for selecting representative MDA's, to judge the reasonable selection of AMDA's, are presented in more detail in the ANSI N13.30 document (24). The considerations of the need to select MDA's that are adequate for individual radiation dose assessments for radiation protection purposes, adequate for detection of the release of unsealed radioactive material from contained processes, and yet not so expensive or time-consuming that they discourage or prevent adequate sampling and monitoring, have been discussed in Section 2.1. In general, the Working Group (24) selected MDA's that are attainable by competent analysts using well-developed procedures and instruments that should, when properly quality controlled, give reliable results. They are not necessarily the best "state-of-the-art" values, in that there may be analytical methods that can provide lower MDA's, but at unreasonable costs for routine bioassay monitoring.

2.7 Considerations for Selection of Acceptable Minimum Detectable Amounts (AMDA's)

Some of the general rationale for selection of AMDA's has been presented in Section 2.1, since the rationale is related to the purpose for establishing AMDAs in the standard (24). Where possible, the AMDA has been established at least several times higher than the MDA, if it is adequate for radiation protection and dosimetry purposes, since then the analytical laboratory could be allowed as much flexibility as possible to use cheaper and more rapid procedures of analysis as available. The latter procedures would again serve radiation protection purposes in the event of emergencies, where rapid turnaround of results may be important. Also, more economical analytical procedures would be important for radiation protection purposes where it would be desirable to monitor greater proportions of a population that might be exposed to many and highly variable sources of potential internal exposure.

For relatively high radiotoxicity nuclides such as plutonium, however, the AMDA may be set closer to the practically attainable MDA, since it serves no radiation protection purpose to discourage operations of laboratories that are performing reliably with the best procedures available, achieving the best MDA's possible (36, 37) for these important nuclides.

Since the determination level of an analytical procedure (7), where detected amounts can be expected to be determined within about a 10% coefficient of variation, may be at least 3 times the MDA, test ranges proposed for the standard (24) generally lie above 3 times the AMDA deemed appropriate for the respective radionuclides and test categories, for purposes of testing laboratory intercomparisons with the bias and precision statistics described in Section 3. However, known spiked test samples may be used near the AMDA level and below to test the variability of s_b or s_D , and thus help ascertain the appropriate determination of MDA for a particular laboratory procedure (perhaps for comparison with the respective AMDA of the standard (24)).

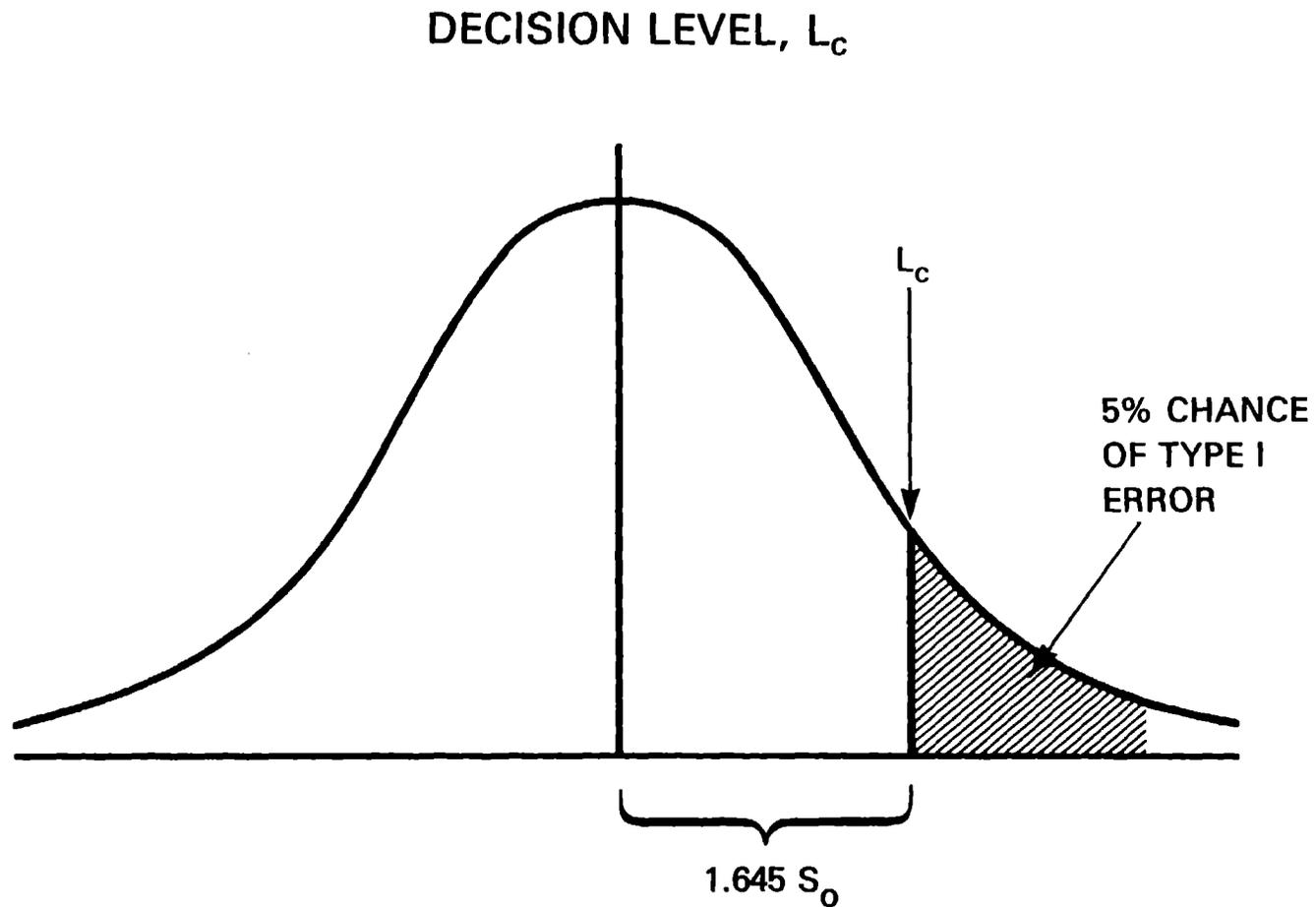


Figure 1 - The decision level, L_C is chosen so that there will be a 5% chance that a net signal from a zero-activity sample compared to a blank will be greater than L_C . Approximately, assuming Normal distribution statistics, $L_C = 1.645 s_0 = 1.645 (s^2_{\text{zero activity}} + s^2_{\text{blank}})^{\frac{1}{2}}$.

DETECTION LIMIT, L_D (PAIRED BLANK)

31

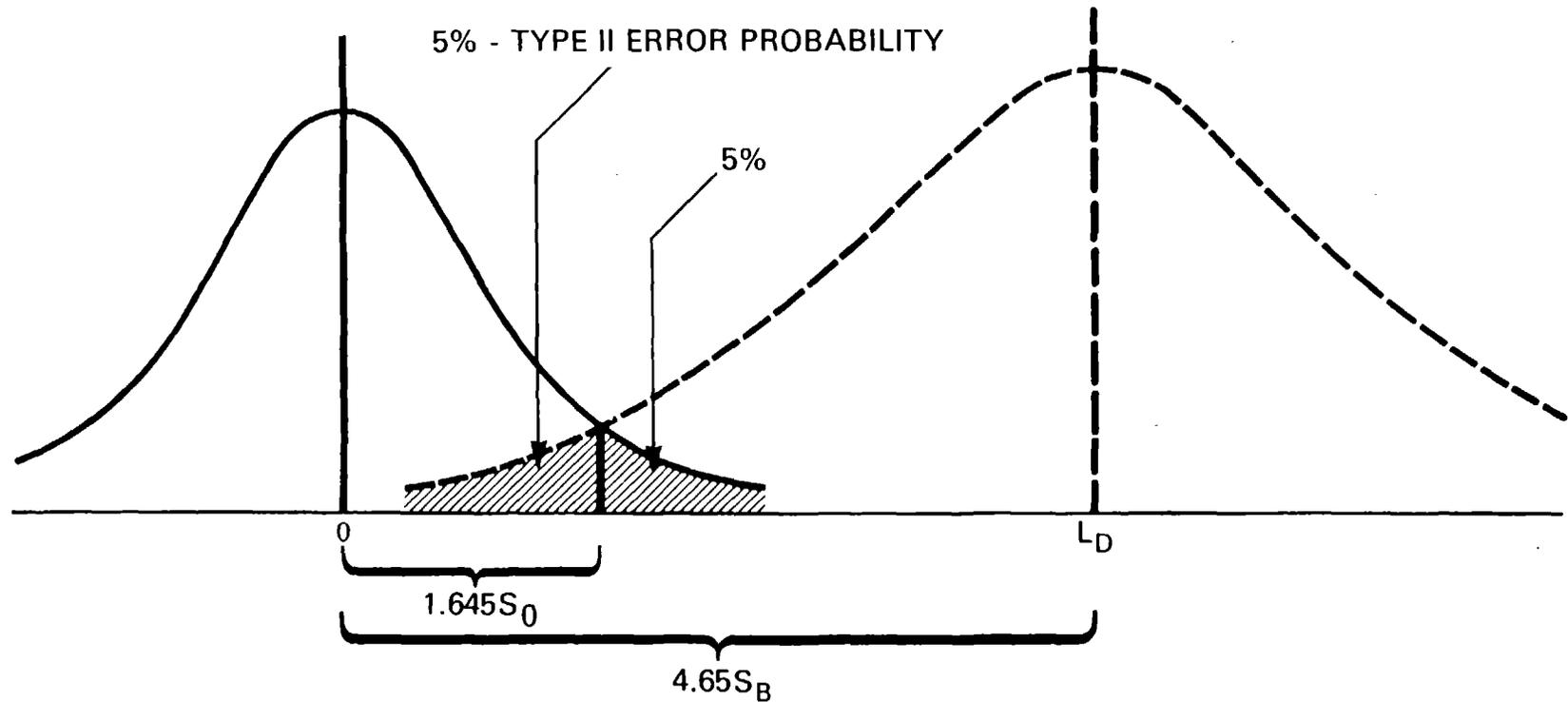
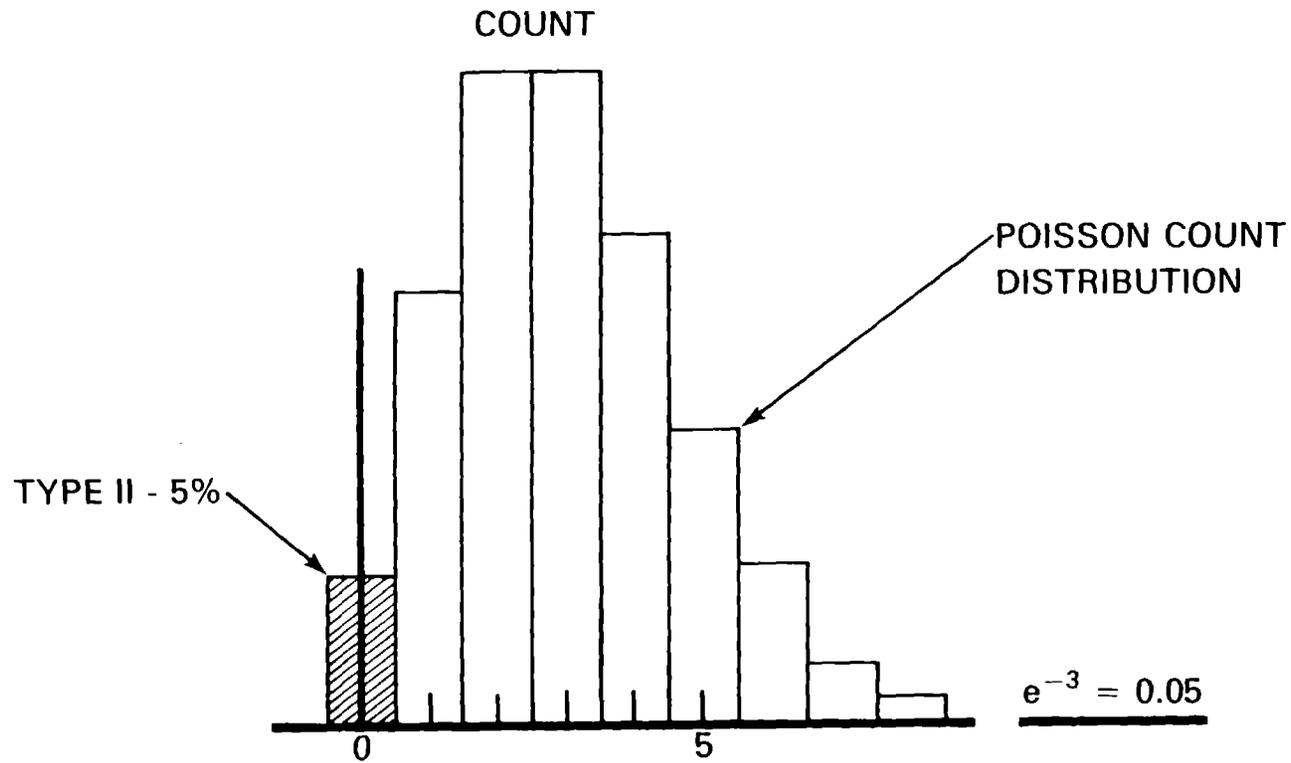


Figure 2 - The detection limit, L_D is placed far enough above zero so that there is a 5% chance that an L_D quantity will give a signal less than L_C . $L_D = 4.65 s_B$ when the counting times of the sample and background are the same, and all s^2 values are approximately the same. Otherwise, $L_D = 1.645 s_0 + 1.645 (s_{LD \text{ activity}}^2 + s_{\text{blank}}^2)^{1/2}$.

ZERO BACKGROUND (BLANK) COUNT



32

Figure 3 - When the number of counts are not large (B greater than about 70), the Normal distribution assumption begins to lose validity; when counts are zero or close to zero, L_C becomes 0, and the detection limit becomes about 3, since $\exp(-3) = 0.05$.

3 RATIONALE AND PRACTICAL IMPLICATIONS OF THE BIAS AND PRECISION STATISTICS, AND THE CRITERIA FOR TEST SAMPLE ACCURACY

3.1. General Approach

Formulations for expressing bias and precision in the standard on bioassay performance criteria (24) are not of obvious origin, and thus the reasons for choosing particular statistics as estimators of bias and precision must be presented here. It is important that the user of this standard understand the purpose in using these statistics as measures of accuracy, and their limitations as well as advantages. Also, the rationale for selecting the acceptable ranges of the bias and precision estimators in the measurement of quality control or test samples will be presented in this Section.

The relative bias estimator $B_{ri} = (A_i - A_{ai})/A_{ai}$ has been selected as the measure of bias for a single measurement so that relative errors can be compared, regardless of units of measurement and magnitudes of the amounts of spike in test samples, with each other between different levels of radioactivity within given test categories, as well as between categories and between different laboratories. Also, in the event that a constant percent bias is inherent within a given category of measurement, within a given laboratory (due, e.g., to a constant error in a calibration source value or a constant relative error in determining a chemical recovery), the estimator B_{ri} would be an "unbiased" estimator of the true deviation (bias) of the mean measurement (the "aim") expressed as a fraction of the actual value as shown in Section 3.2 below. An unbiased estimator B_{ri} , in statistical terminology, is a "best" estimator in the sense that it would in the long run* provide average values that would converge on the true underlying relative bias, and perhaps provide clues to the removal of such bias and the improvement of the accuracy and reliability of analyses.

Recognizing that no small sample of test measurements A_i , relative to respective known spiked quantities designated as A_{ai} , can determine a measurement bias independent of random fluctuations of measurement, the draft ANSI standard (24) calls for enough measured values A_i in a given test category (within economic limits) so that an average value B_r may be calculated that will to some degree "average-out" some of the fluctuations of measurement and serve as an approximate measure of the underlying relative bias in a given category. Then, one measure of precision (dispersion) of a given category is taken as an estimate S_B of the variation of individual B_{ri} values about the average value B_r . Another measure of precision, the coefficient of variation, S_A , of the sample measurements about their average, is also used for reasons discussed in Section 3.5. The precision estimator S_B corresponds to a variance S_B^2 that is an unbiased estimator of the true variance of B_{ri} , which is the same as the true variance of the sample measurements divided by the square of the true (spiked) value (see Section 3.3).

*"...in the long run..." means that as more data are collected and incorporated into the calculations of bias and precision statistics, under constant procedures for quality control testing.

The statistic S_B is also stochastically independent of B_r when the random errors of measurement occur in factors that are additive in formulating results (22, 23). This independence and unbiasedness provide an advantage that B_r and S_B can provide long-range estimates of any fixed errors separate from the random errors of measurement. S_B multiplied by the true (spike) value will also be an estimate of the true standard deviation of measurement. After calculation from the data of an inter- or intralaboratory test run,) the estimators of bias, B_r , and precision, S_B and S_A , are then compared, for each spike value and each measurement category, to the acceptable ranges of error given in the draft standard (22, 24).

The acceptable ranges of error are chosen by the Working Group of ANSI N13.30 (24), based on their cumulative analytical experience together with suggestions from advisors and commenters on various drafts of the standard. Also, results of previous interlaboratory tests, the relative accuracy needs of bioassay data for estimation of internal radiation doses to workers and members of the public, and the probabilities that adequate state-of-the-art laboratories could pass reasonable tests, have been taken into account in selecting acceptable accuracy limits.

Most of the test samples, in the application of this standard, should be designated to be well above the Acceptable Minimum Detectable Amounts (AMDA's) of each radionuclide in each test category, as indicated by the test ranges specified in the draft (or final) standard (24). Then, the unavoidable fluctuations of measurement such as the random Poisson variations in a radioactivity count (which would not have a constant relative deviation with respect to the amount of activity measured) can be kept relatively small compared to any other sources of bias or dispersion error (which may hopefully be determinable and correctable). Currie (7, 20) has suggested that a measurement should not be considered to be an analytical determination unless it is above a "determination limit," which is in effect at least about 3 times the "detection limit" (corresponding to a "minimum detectable amount" (MDA), as discussed in Section 2). (See Currie (7), Tables I and II). Also, even with considerable replication, in low level measurements bias may go undetected until it exceeds the standard error of measurement by a factor of four or more, (See Currie, Gerlach, Klouda, Ruegg and Tompkins (47), p. 554). Thus, the measurement range over which bias and precision are to be tested based on the statistics B_r , S_B and S_A calculated from small samples should be at least above 3 AMDA. For some laboratories, the MDA may actually be at or near the AMDA. For some radionuclide measurements (e.g., ^{239}Pu in lung), the AMDA has been selected for this standard close to the "best" attainable MDA, since the relative radiotoxicity and radioactive emissions of this radionuclide are such that the ultimate is required in measurement capabilities for radiation protection purposes (48-49).

3.2. Demonstration of the Unbiasedness of B_{ri}

Consider a sample space of measurements A_i obtained from test samples containing relatively exact (assumed constant) amounts A_{ai} of radioactivity of a certain nuclide. Assume that there is a bias error b in the measurement system, so that the expected value $E(A_i)$ is

$$\text{defined } E(A_i) = \int_{A_i = -\infty}^{A_i = +\infty} A_i p(A_i) dA_i = A_{ai} + b \quad (\text{Eq 60})$$

where $p(A_i)$ is the probability density function of the continuous variate A_i , and b is the (+ or -) bias error.

For a very large number of measurements under constant conditions, there would be a sampling distribution of values of B_{ri} calculated from measurements A_i . The expected value of B_{ri} would be

$$E(B_{ri}) = E\left(\frac{A_i - A_{ai}}{A_{ai}}\right) = \frac{1}{A_{ai}} \int_{-\infty}^{+\infty} (A_i - A_{ai}) p(A_i) dA_i, \quad (\text{Eq 61})$$

since $p(A_i - A_{ai}) = p(A_i)$ and $d(A_i - A_{ai}) = dA_i$, there being a 1:1 correspondence in event space between $A_i - A_{ai}$ and A_i .

Using the distributive properties of integration,

$$\begin{aligned} E(B_{ri}) &= \frac{1}{A_{ai}} \int A_i p(A_i) dA_i - \frac{A_{ai}}{A_{ai}} \int p(A_i) dA_i \\ &= \frac{1}{A_{ai}} E(A_i) - 1 = \frac{1}{A_{ai}} (A_{ai} + b) - 1 = 1 + \frac{b}{A_{ai}} - 1 \\ &= \frac{b}{A_{ai}}, \end{aligned} \quad (\text{Eq 62})$$

which is the underlying relative bias for measurements A_i .

Thus, the average of B_{ri} , B_r , will converge even more rapidly to b/A_{ai} , for a particular sample spike value A_{ai} , as sample size N increases, since

$$B_r = \frac{1}{N} \sum B_{ri}, \text{ and} \quad (\text{Eq 63})$$

$$E(B_r) = \frac{1}{N} \sum E(B_{ri}) = \frac{1}{N} \cdot N \frac{b}{A_{ai}} = \frac{b}{A_{ai}} \quad (\text{Eq 64})$$

3.3 The Unbiasedness of the Precision Measure S_B^2

Using the selected formula for S_B given in the draft standard (24).

$$E(S_B^2) = E\left(\sum_{i=1}^N (B_{ri} - \bar{B}_{ri})^2 / (N-1)\right), \quad (\text{Eq 65})$$

and we want to prove that $E(S_B^2)$ is identical with the variance $\sigma^2(B_{ri})$ in measurements B_{ri} about $E(B_{ri}) = b/A_{ai}$,

where

$$\begin{aligned} \sigma^2(B_{ri}) &= E\left(B_{ri} - \frac{b}{A_{ai}}\right)^2 = E(B_{ri}^2) - E(B_{ri})^2 \\ &= E(B_{ri}^2) - \frac{b^2}{A_{ai}^2}. \end{aligned} \quad (\text{Eq 66})$$

Expanding the formula for $E(S_B^2)$,

$$E(S_B^2) = E\left(\sum_{i=1}^N (B_{ri}^2 - 2B_{ri}\bar{B}_{ri} + \bar{B}_{ri}^2) / (N-1)\right) \quad (\text{Eq 67})$$

$$= \frac{1}{N-1} \sum_{i=1}^N (E(B_{ri}^2) - E(2B_{ri}\bar{B}_{ri}) + E(\bar{B}_{ri}^2)) \quad (\text{Eq 68})$$

Now,

$$\sum_{i=1}^N E(B_{ri}^2) = N E(B_{ri}^2), \quad (\text{Eq. 69})$$

Because $E(B_{ri}^2)$ is a constant value for a given measurement quantity A_{ai} in the spike. Also,

$$\begin{aligned} E(B_{ri} \cdot 2\bar{B}_{ri}) &= 2E\left(B_{ri} \cdot \frac{\sum_{j=1}^N B_{rj}}{N}\right) \\ &= \frac{2}{N} (E(B_{ri}^2) + \sum_{j \neq i}^N E(B_{ri}) \cdot E(B_{rj})) \end{aligned} \quad (\text{Eq 70})$$

since $E(X \cdot Y) = E(X) E(Y)$ when X and Y are statistically independent. Equation (70) then becomes

$$E(2 B_{ri} \cdot \bar{B}_{ri}) = \frac{2}{N} (E(B_{ri}^2) + (N-1) \frac{b^2}{A_a^2}) \quad (\text{Eq 71})$$

since both $E(B_{ri})$ and $E(B_{rj}) = b/A_{ai}$, being sample variates of the same frequency distribution. The third term in brackets in Equation (68) becomes:

$$E(\bar{B}_{ri}^2) = E\left(\left(\sum_{j=1}^N B_{rj}/N\right)^2\right) = \frac{1}{N^2} E\left(\left(\sum_{j=1}^N B_{rj}\right)^2\right) \quad (\text{Eq 72})$$

$$= \frac{1}{N^2} E\left(\sum_{i=1}^N (B_{ri} \left(\sum_{j=1}^N B_{rj}\right))\right) = \frac{1}{N^2} (N E(B_{ri}^2) + N(N-1) E(B_{ri}) E(B_{rj}))$$

$i \neq j$

$$= \frac{1}{N} E(B_{ri}^2) + \frac{N(N-1)}{N^2} \frac{b^2}{A_a^2} \quad (\text{Eq 73})$$

Substituting the relationships from Equations (69) through (73) into Equation (68), we get:

$$E(S_B^2) = \frac{1}{N-1} (N E(B_{ri}^2) - 2 E(B_{ri}^2) - 2(N-1) \frac{b^2}{A_a^2} + E(B_{ri}^2) + (N-1) \frac{b^2}{A_a^2}), \quad (\text{Eq 74})$$

which upon cancelling and combining appropriate terms becomes

$$\begin{aligned} &= \frac{1}{N-1} (N E(B_{ri}^2) - E(B_{ri}^2) - (N-1) \frac{b^2}{A_a^2}) \\ &= E(B_{ri}^2) - \frac{b^2}{A_a^2} = E(B_{ri}^2) - (E(B_{ri}))^2, \text{ which is} \\ &= \sigma^2(B_{ri}) - \text{the variance of } B_{ri}. \end{aligned} \quad (\text{Eq 75})$$

Thus, the expected value of the relative variance statistic S_B^2 (the square of the defined relative precision estimator is the "true" underlying population variance $\sigma^2(B_{ri})$, as desired.

3.4 Other Desired Characteristics of the Precision Statistic, S_B

There are other desirable characteristics of the precision measure S_B in addition to the unbiasedness of its associated variance (S_B^2). As defined in this standard, S_B^2 can be shown to be algebraically equivalent to the variance estimator s_A^2/A_a^2 , and thus $S_B = s_A/A_a$, where s_A is the standard deviation of sample measurements of A at some given true spike level A_a in a given measurement category.

That is, $s_A = \sqrt{\sum (A_i - A_{ai})^2 / (N-1)}$, with terms as defined in Section 3.2, and A_{ai} is the actual (true) value of the spike for the i^{th} replicate of a sample, explicitly or implicitly traceable to a National Bureau of Standards measurement within usually a few percent. When all replicate test samples in a given test measurement are spiked at the same level, then all $A_{ai} = A_a$.

Thus, the precision estimator S_B as defined for use in this standard has the following additional desirable characteristics: (1) the absolute value of the dispersion of a replicated measurement, s_A , may be estimated simply by multiplying S_B by the true spike value A_a , for use in MDA calculations or in examining the absolute dispersion of measurements for purposes on internal quality control; (2) the estimators S_B and s_A are both independent of the estimator B_r or the true bias b when the random errors between the i^{th} samples are additive (50), and thus are not confounded by changes in the bias due to changes in "systematic", "fixed", or "deterministic" errors; and (3) thus the dispersion in measurement of a quantity can be examined consistently from one series of measurements to another, whether or not the true bias (alone) of the underlying system has been changed due to changes in experimental procedures or calibration factor errors.

3.5 Characteristics of the Precision Statistic, S_A , the Coefficient of Variation

The coefficient of variation of a measurement, $S_A = s_A / \bar{A}$, where \bar{A} is the average of the measured values in a series of replications or measurements, is often of interest and useful for internal quality control purposes, and is a familiar statistic to all scientists. Also, when the random errors between replicates are "multiplicative" (i.e., when the random errors are in terms such as the counting efficiency E , the sample size V , or the yield Y in Equation 1, or the overall "calibration factor" K of Equation 50, rather than the error Δ in Equation 50, then S_A , rather than S_B , is independent of the bias estimator B_r (22, 23, 50). The value of \bar{A} is, of course, dependent on the bias of the associated measurement. However, a multiplicative error cancels in the ratio $s_A / \bar{A} = S_A$. Since the limited studies so far indicate that both "additive" and "multiplicative" random errors may be important in the various analytical categories, both S_A and S_B are proposed for use in the draft standard so that both can be examined (22-24). Whenever the coefficient of variation is recorded for internal quality control purposes, the associated values of \bar{A} and A_a should be maintained as part of the same record so that S_B as well as S_A may be determined for the purposes of examining ongoing quality control or intercomparison programs, and for calculating any relevant probabilities of passing any postulated or real precision criteria.

The above rationale for the use of S_A and S_B as precision estimators, together with the use of B_r as a bias estimator, is applicable to the establishment of consistent methods of assessing quantitative accuracy performance, regardless of whether the random fluctuations of measurement values A_i are distributed normally, lognormally, or according to some other skewed frequency distribution. The choice of estimators of bias and precision is somewhat arbitrary, but

several desirable characteristics (Reference 5, page 37) have been considered: unbiasedness; minimum variance; convenience in calculations; and ease in calculating probabilities and implications of selected performance criteria for the bias and precision estimators. Later sections of this report illustrate the utility of probability calculations using the selected estimators B_r and S_B or S_A . The calculations of S_B are easily carried out directly from the B_{ri} values utilized in calculating B_r . S_A is simply the coefficient of variation of replicate measurements. Thus, the estimators B_r , S_B , and S_A provide convenience and economy of calculation, in addition to providing otherwise desirable statistical estimates of relative bias and relative precision.

3.6 Rationale for the Limits on Bias and Precision

3.6.1 Specific Accuracy Criteria Proposed

The ranges of acceptable bias and precision in the standard (24), over the quantity ranges (times AMDA) selected, are specified in terms that allow objective decisions of acceptability based on the specific estimators of B_r , S_B , and S_A , calculated from small sample data. The requirements that B_r be within -0.25 to 0.50 , and that both S_B and S_A be in absolute value less than 0.4 (sample bias estimate within -25% to 50% , and sample precision within 40% , respectively) actually require that a laboratory maintain accuracies of about 10 to 15 percent in underlying bias, and about 10 to 20 percent in relative precision, in order to have a reasonably high probability of repeatedly passing the criteria for accuracy in the quantity ranges judged important and feasible for analytical determinations (7) for bioassay purposes. (See Table 4). When the underlying precision is such that the standard deviation σ of the logarithms of A_i for lognormally distributed data is 0.1 (10%), the coefficient of variation for the lognormal data would be $\{\exp(\sigma^2) - 1\}^{1/2}$, or only 0.1003 (see pages 8 and 154, Reference 51). Then, the standard geometric deviation S_g would be $\exp(0.1) = 1.1052$, and the 68 percent range of the underlying distribution would be $\bar{x} \pm 1.1052$, or 0.904 to 1.1052 (52). This is close to the 0.90 to 1.10 range for 68 percent of the data for a normally distributed population. Thus, lognormally distributed data would be approximately normal for measurements within the acceptable range of precision. Also, for similar reasons, values of S_A will not be very different from those of S_B when the coefficients of variation in measurements are no greater than 10 - 15% (50).

While much of the data and error distributions of radiation protection measurements have been found to be lognormal (52, 53), tests of the radioactivity determinations in bioassay and other radiochemical laboratories indicate that the observed data are often more likely to approximate a normal distribution much of the time (22, 23, 54). This is probably true because, for low level radioactivity measurements from biological samples, the subtraction of background-plus-appropriate blank measurements from respective sample-plus-blank-plus background measurements will allow some results at very low levels to fall into the negative data range in such a manner that the data become very closely approximated in distribution by the normal probability density function. Thus, the example calculations in later sections of this report are carried out for

normally distributed data, although methods and algorithms are also presented for use with lognormally distributed data. These example probability calculations are provided to illustrate the pass-fail implications of establishing various quantitative performance criteria of bias and precision in an analytical performance standard.

3.6.2 General Considerations in Establishing Standards of Accuracy

The selection of standards of accuracy for bioassay measurements, as embodied in the objective ranges of acceptability for the sample statistics B_r , S_B and S_A require the balancing of a number of considerations that often compete with each other. These considerations, deliberated in many sessions by the working group (24) in a critical examination of the accuracies needed for the various bioassay measurements, fall into several categories: those considerations of accuracy needed in order to provide optimal data for detecting and measuring internal human burdens of radioactive material, for ultimately estimating doses and dose commitments for risk estimation and radiation protection management purposes and, where appropriate, for the medical management of chelation therapy; the considerations for providing bioassay and internal dose data suitable for meeting existing legal and record-keeping requirements; the "state-of-the-art" of the various radioanalytical in vitro and in vivo methods for determining quantities and concentrations of radionuclides in representative media of primary interest for radiation protection purposes (46); and the need to select accuracy standards that, while meeting basic radiation protection needs, would be reasonable and economically feasible for routine commercial or private laboratory service, so that the costs of sample measurements do not limit unduly their adequate use for radiation protection purposes as recommended by other existing standards of bioassay program management. These several areas of consideration will be reviewed in the paragraphs below. Previous published guidance on accuracy requirements will be summarized in Section 3.6.3. Some of the rationale developed for balancing these considerations to arrive at specific accuracy standards will be summarized in Section 3.6.4.

3.6.3 Summary of the Literature and Considerations Regarding Accuracy Requirements

3.6.3.1 Radiation Protection Accuracy Requirements

In the routine monitoring of persons for radiation protection purposes, procedures must be established to ensure that workers have exposures measured and recorded with a reasonable degree of accuracy so that their exposures can be controlled and maintained ALARA, and so that they will not inadvertently during the course of their employment exceed the recommended and regulatory limits of quarterly or annual exposure. "Action" or "investigation" levels are often established that are based usually on indirect measurements that indicate exposures may be occurring for which investigations or measures should be undertaken to change working conditions, or increase surveillance to better control what appears to be an increasing exposure. The ICRP in 1968 (55) suggested that acceptable uncertainties of measurement could be related to these "investigation levels" and stated, "The uncertainties acceptable in routine individual monitoring should be somewhat less than the investigation level and can best be expressed in relation to the annual dose. The uncertainty in assessing the upper limits to the annual dose equivalent to the whole body or to the organs of the

body should not exceed 50%. Wherever these doses are less than 2 rem, an uncertainty of 1 rem is acceptable." This recommendation is not accompanied by any rigorous derivation or supporting information, but seems to be the consensus of knowledgeable persons regarding the need for some confidence that dose limits will not be exceeded while taking into account the inherent variability of single measurements of radiation dose and single estimates of radionuclide intakes. The meaning of "uncertainty" is not well defined.

The National Council on Radiation Protection and Measurements (NCRP) (56) has called for an "accuracy" of ± 30 percent at doses approaching maximum permissible levels (MPD) and ± 20 percent at higher doses. The NCRP indicates, however, that a factor of 2 uncertainty (which would mean for radiation protection purposes a range of -50% to +100%) would be acceptable at doses below 1/4 MPD (see pp 63-64, Reference 56). The same report calls for a precision of $\pm 10\%$ in personnel monitoring to improve comparisons of trends between persons and time periods. Internal exposure monitoring is recommended when air concentrations may possibly lead to depositions of 10% of MPOB (p. 69, Reference 56) but an accuracy of $\pm 30\%$ is suggested for activity or dose estimates from internal emitters (p. 71, Reference 56)). Again, "accuracy", "uncertainty", and "precision" are not rigorously defined or derived, but are apparently left to the judgement of the individual radiation protection manager.

In terms of estimation of individual risks for risk management purposes, it would be difficult to justify very exacting accuracy requirements in general terms. Uncertainties in the point estimation of risk factors for given doses below permissible limits may be factors of 2 to 10 (57). Even if the point risk estimates were known for every dose level, (and for every type of radiation, exposure rate and portion of body exposed), the individual biological variations in responses of different animals, even from a homogeneous species and strain, may vary within a factor of 2 or more even for the so-called "non-stochastic" effects (58) that are observable at higher dose levels. For example, dose-response shapes tend to approximate log-normal functions, and the standard geometric deviation of the 30-day mortality curve for external gamma-ray irradiation of mice (Figures 17 and 18, Reference 59) can be seen to be about $s_g = \frac{x}{\div} 1.2$. The 95 percent range of variability of the lognormal dose-response distribution would be about $(1.2)^2 = \frac{x}{\div} 1.44$; i.e., 95 percent of animals would die at doses within a factor of 1.44 of the median lethal dose or between 500 to 1050 rads. A similar log-normal variability was found for lethality from high-energy protons of two differing distributions of linear energy transfer (60). Interpreting the variability of the dose required for a quantal response of death as a biological variability in animal sensitivity (61), this means that there is a factor of 2 difference between the lower and upper 95 percent range limits on the sensitivities of individual animals in these experiments, which pertain to homogeneous mammals uniformly maintained and accurately exposed under laboratory conditions.

The variability of response of animals to "stochastic" effects such as cancer at low dose levels of external radiation also shows factors of two or more in dose between the 5% and 95% response levels. At least some of this variation may be attributable to the stochastic nature of the response rather than a true inherent biological variability between animals (53, 62).

For irradiation of tissues by radioactive materials dispersed or implanted within the body, the variability of response is even greater than for external irradiation. Both for alpha and for beta radiation, the ranges of dose-response curves often extend over two orders of magnitude (some individuals die at 100 times the dose that kills others by malignancy), both for animals (62) and for humans (63). Over a wide range of internal dose, the incidence of carcinogenesis from radium in bone was not very dependent upon the exact dose, once doses exceeded about 1000 rads to bone (Fig. 3, Reference 63). Since the exact positions and shapes of the dose-response functions are not known even to within factors of 2 to 10 in response magnitude, it is evident that small errors in dose - even a factor of two - would not appreciably affect the order of magnitude estimates of risk that are made for protection design purposes.

In radiation protection design and operations management, factors of safety of 100 often are necessary to protect adequately against internal radiation exposure (64). These large safety factors may often nullify any effect on absolute risk of errors in dose estimation, if in fact the doses that are recorded give a proper indication of trends and are utilized to effectively maintain exposures ALARA. For example, if an indication of the possibility of appreciable internal exposure is obtained from bioassay samples in a particular instance, the employer may either find it reasonably inexpensive or estimate it cost-effective, to move the particular operation of concern to a ventilated hood - in which case the further exposures will most likely be reduced by at least factors of 10 to 100. If for similar reasons an operation is moved into a completely enclosed properly maintained glovebox, the reduction in future internal exposure to the worker may typically be as much as a factor of 100 million (64). Thus, uncertainties of as much as a factor of two in dose or dose commitment estimates may not seriously influence actual risks to workers, or the radiation protection equipment and procedures provided to them, as long as the safety factors customarily used in radiation protection planning and operations are applied.

In deriving standards of accuracy for bioassay, it is also important to consider that an individual bioassay sample determination is usually used only to determine whether additional actions of follow up are needed, and not as a parameter by itself for calculating internal doses or ultimate dose commitments (65). A wider range of variability in the measurements due to random errors is allowable than would be allowed for the final dose estimation, if we can assume most of the error is due to inherent random variations. This is true because, if the calibrations are proper and the bias of measurement is adequately small, then the statistical laws of large numbers will ensure that the ultimate estimation of intake will have narrower limits of precision than any single measurement (66). The standard deviation S of N measurements X_i would of course become a standard error $S/(N)^{1/2}$ of the average X used as an estimate of the underlying quantity, if the quantity were constant for all measurements. For a retention function estimated from M data points (or averages of replicates) taken over time the variance of the estimated retention at each sampling time is approximately inversely proportional to M , so that the relative coefficient of variation will decrease approximately as $1/(M)^{1/2}$ (p. 142, Ref. 67). (In order to obtain as accurate an intercept at $t = 0$ as possible, and better definition of the retention function at early times after a known intake, it is better to take frequent measurements at early times, perhaps daily excretion samples and *in vivo* counts for the first three weeks, in order to obtain a more accurate definition of the applicable metabolic models and a check on the initial distributions of the inhaled material to determine which ICRP lung and GI tract model parameters are

applicable.) Thus, if the relative precision S/μ is 40 percent for a single bioassay determination, due mainly to random errors fluctuating about a relatively accurate value, then a triplicate determination would have a relative standard error of $40/(3)^{\frac{1}{2}} = 23.1\%$. Then, if 17 urine samples were analyzed over a period of weeks, as in the Pu-Am inhalation case described in Reference 68, parameters estimated for an appropriate retention model might be expected to yield relative errors of only about $23.1/(17)^{\frac{1}{2}} = 5.6\%$ in estimating the retention at a particular point in time. In any realistic situation, an error of about 6% would be only a small proportion of the likely systematic error due to uncertain calibration factors, low chemical recoveries, or many other sources of error in bioassay determinations (46, 69, 70).

In the same Pu-Am case, estimates of lung burden at given points in time often varied by factors of two or more between measurements, except for the effects of spurious surface contamination that was detected and removed on day 5 (Table 2, Ref. 68). The random fluctuations by a factor of two were primarily a result not of Poisson variations in counting statistics, but were due to fluctuations in chest wall and rib attenuation of the Pu-239 x-rays and Am-241 60-keV gamma-rays at different measurement positions on the chest. Differences of almost a factor of two were observed between systemic burden estimates from urinalysis and body burden estimates from in vivo counting, for a person exposed to a mixture of ^{241}Am and ^{239}Pu (about 9:1 activity ratio in air sample) after a glove-box explosion of an acid Pu-Am solution (Reference 69, Table V). Systematic uncertainties in these estimates included those due to: initial external widespread contamination; differences in Am/Pu ratios between air samples and fecal samples (Am/Pu ratio slightly higher in feces); uncertainties in the effects of chelation treatments on excretion rates; and uncertainties in calibration factors using simulated phantoms for in vivo determinations. Such systematic uncertainties are common to all laboratories, and provide dose estimates from the long-lived, more radiotoxic materials, that can not usually be assumed reliable to better than a factor of two (65, 68, 69, 70).

Snyder (70) has shown that by appropriately allowing for the statistical variability in urinary output of plutonium, it was possible to estimate the internal burdens of two individuals to within a factor of two of their burdens as obtained from autopsy samples. However, his method of computation requires at least 20 sample urine measurements in order to adequately "average out" the random fluctuations in daily urinary output of plutonium. More recent autopsy results (71) show that, although re-analysis of urinary data by six of the best laboratories shows usually good agreement between the original estimates of body burden by health physicists cognizant to the original investigations, the estimates of burden from long-term urinary data were a factor of 1 to 22 (averaging about 4.5) times higher than the burdens estimated from the direct tissue analyses of plutonium in samples taken at autopsy. Although part of this overestimation might be due to the limited sizes of tissue samples at autopsy (71), nevertheless the consistency of the overestimation indicates that, over the decades, the health physics profession has deemed it necessary to employ conservative methods of internal burden and dose estimation to allow for the large uncertainties in converting bioassay data into estimates of intake, burden, and dose or dose commitment.

Thus, a dilemma is created in rationally specifying "acceptable" accuracy on the basis of radiation protection needs. Both narrow and wide limits of accuracy

have been recommended by national and international expert committees, with little rationale or consistency between recommendations. Obviously, uncertainties of as much as a factor of 2 in dose determinations may not be as important in risk assessment as the uncertainties in risk per unit dose, nor may such uncertainties influence the decision to provide additional protective measures in the form of facilities, equipment or radiation safety procedures, as long as sufficient safety factors are already incorporated into the operation and trends are appropriately indicated by the monitoring systems. On the other hand, records of radiation exposure may all too frequently give erroneous indications of overexposure, or serious underestimates of exposure, when dosimetric accuracy limits are too wide. Narrower limits of error would be more consistent with the practical needs of radiation protection program management, and certain legal and regulatory considerations, which will be reviewed briefly in the following sections.

3.6.3.2 Administrative Considerations

There are no specific quantitative criteria presented in previous American National Standards (72, 73) that can be used to establish an administrative requirement for the accuracy of bioassay measurements or internal dose estimates, although for tritium bioassay, the Appendix C of a very recent standard (Reference 74, p. 15) provides an adaptation of ICRP recommendations (75) to make suggestions regarding the accuracy required in estimating the "... upper limit to the actual dose equivalent that could have been received or committed. These criteria are that

- (1) The estimates over a year should be within 1 rem of the upper limit (at the 95% confidence level) if the upper limit is not greater than 2 rem in the year.
- (2) The estimates over a year should be within 50% of the upper limit (at the 95% confidence level), if the upper limit is greater than 2 rem in the year."

This statement of accuracy requirements adds what appears to be a 95% confidence interval interpretation to the ICRP statements (75) that, "The uncertainty in assessing the upper limits to the annual dose equivalent to the whole body or to the organs of the body ... should not exceed 50 percent. Where these doses are less than 2 rems an uncertainty of 1 rem is acceptable. This uncertainty includes errors due to variations in the dosimeter sensitivity with incident energy and direction of incidence, as well as intrinsic errors in the dosimeter and its calibration." Thus, the ICRP statement of accuracy provides a judgmental view of the degree of variability of field conditions influencing dose determinations, but does not provide a precise definition of accuracy required of the dosimeter itself, the basic initial quantity to be determined before interpretation.

Since development of the tritium bioassay standard (74), the ICRP has issued ICRP Publication 35 (76), revising its recommendations of ICRP 12 (75) and its statements regarding accuracy requirements. In addition to recommending that uncertainties of measurement of annual deep and shallow dose-equivalent received from external exposure be "reduced as far as reasonably achievable," the following statements are made: "If these quantities are of the order of the relevant annual limits, the uncertainties should not exceed a factor of 1.5 at the 95%

confidence level. Where they amount to less than 10 mSv an uncertainty of a factor of 2 at the 95% confidence level is acceptable. This uncertainty includes errors due to variations in the dosimeter sensitivity with incident energy and direction of incidence, as well as intrinsic errors in the dosimeter and its calibration. It does not include uncertainties in deriving tissue or organ dose equivalents from the dosimeter results." (Reference 76, page 25.)

In the revised recommendations for monitoring workers (76), the ICRP also includes a statement in regard to accuracy requirements in routine monitoring for internal contamination:

"Ideally, the uncertainties in routine individual monitoring for internal contamination should be similar to those accepted in monitoring for external contamination ... In practice, however, uncertainties as small as 50% are rarely possible in routine monitoring for internal contamination because of the errors introduced by the lack of knowledge about the pattern of intake and retention, combined with the errors in measurement. In these circumstances it is not appropriate to set a defined objective in general terms. In making use of the measurements, and particularly in informing the individual concerned about intakes or committed dose equivalents, the magnitude of the uncertainties should not be overlooked." This is probably the most recent statement regarding accuracy requirements of internal dose estimates as stated by an authoritative radiation protection committee.

The ANSI standard for internal dosimetry for fission and activation products (77) does not give specific accuracy criteria for single measurements of either in vitro or in vivo bioassay samples, but does call for considerable quality assurance procedures, and the determination of precision by replicate measurements, and states that "... every effort should be made to reduce systematic (or non-random) errors that could affect results" for in vivo measurements. Also, the in vivo facility criteria call for designs that would "... allow measurements of 5% of the MPOB of the radionuclides listed in 8.2 for at least 95% of the in vivo measurements performed." (Reference 77, page 18). This latter quotation, together with statements regarding the need to determine precision by including all sources of random error in the replicate measurements, could be used as a basis for determining acceptable minimum detectable amounts (AMDA's) of the radionuclides included in the ANSI N343 standard. The determination of acceptable MDA's would also imply a required absolute precision at the AMDA level.

For example, if s is the absolute precision of a measurement of the net amount in microcuries of Am-241 in the liver (using, e.g., a collimated detector), and if s_b is the standard deviation in measuring an unknown blank phantom containing the appropriate amount of K-40 but no Am-241, then at a decision level of $2.33 s_b$ a determination would be called positive if the probability is set at the 0.05 level for a Type I error (20).

The MDA for a probability of 0.05 of a Type II error would be about $2.33 s_b + 1.645 S$. If a total number of counts greater than about 70 is taken for the measurement, then the MDA for a 0.05 probability of a Type II error would be about $4.65 s_b$, assuming $s = \sqrt{2} s_b$ (see Section 2). Interpreting ANSI N343* as

*The working group for this standard is being reconvened in January 1986 to consider whether this standard should be revised to incorporate more recent national and international standards of dose limitation and radiation protection.

intending that $4.65 s_b = 0.05 \text{ MPOB}$, then the AMDA is inferred to be 0.05 MPOB, and the desired acceptable underlying precision of measurement at the MDA level is $s_b = 0.05 \text{ MPOB}/4.65 = 0.0108 \text{ MPOB}$ (about 1% of a "maximum permissible organ burden" (32)). The relative underlying precision of the population of measurements is therefore $s_b / 4.65 s_b = 21.5\%$ at the AMDA level. If the random errors of measurement are primarily the Poisson counting fluctuations, then at 3 AMDA, the relative reduction in error will be about $(3C)^{1/2}/(3C)$, so it will be somewhat reduced depending on the total count C. If the random errors are mostly due to fluctuations in the geometry of phantom placement or other variables that tend to have a constant proportional error, then the relative error may not be reduced appreciably as the level of radioactivity is increased. If the random variations are due primarily to fluctuations in shielding from the detector a constant quantity of interfering contamination in the counting chamber, then the relative precision may improve considerably as the amount of Am-241 in the phantom (or person) is increased. Thus, the dependence of relative precision on amount of radioactive material present above the MDA can be seen to depend critically on all of the specific circumstances of the measurement process. Thus, the suggestion (77) that *in vivo* measurements be able to measure 5% of an MPOB does not provide any guidance on overall accuracy desirable for bioassay measurements in general, but may be taken to indicate that a precision of 1% of an MPOB is desirable where attainable.

International Atomic Energy Agency safety standards (78) indicated that an accuracy of $\pm 50\%$ in external radiation personnel monitoring is adequate for radiological protection purposes, although better accuracy is often desirable when

"... readings from instruments with different specifications are to be intercompared." (Reference 78, page 8.) However, in the assessment of dose equivalent, the same standard points out that dose distribution within the body and other factors are usually unknown so that the "...assigned dose equivalent to the whole body may be regarded as the upper limit." (Reference 78, page 12.) No general statement of accuracy requirements for internal dose assessment are given in this standard, but it is recognized that accuracy requirements are interdependent with investigation levels and measurement frequencies, and that additional data beyond the initial samples will be needed to improve accuracy of internal dose assessments and evaluate the extent of uncertainties (Reference 78, page 15).

3.6.3.3 Accuracy Requirements for Bioassay as Implied by Uncertainties in Air Monitoring

Another IAEA standard (79) recognizes that while air monitoring for radioactive material concentrations is an important part of surveillance of the work environment in many circumstances, and may be the only means of assessing approximate individual exposures in many instances, that uncertainties in the parameters needed to relate air concentrations to internal radiation exposure "...are so large that other methods such as bioassay and whole-body counting are generally preferable," for purposes of assessing actual doses to individuals (Reference 79, page 4). Yet, while recognizing that different "sensitivities" are required for different air sampling purposes and circumstances, the same standard suggests that it is generally desirable to measure concentrations at least to 1/10 MPC, or in exceptional cases to 10 or 1000 times MPC. The standard defines

sensitivity by the statement, "Sensitivity may be expressed as the lowest concentration that can be measured with a precision of $\pm 50\%$ or as a multiple of the background reading." (Reference 79, page 12.)

From this statement it may be inferred that the authors of the standard are referring to the quantity that is the minimum detectable amount (MDA) corresponding to the limit of detection, L_D (1, 7), in units of concentration. Then, the standard implies that the MDA of concentration must be related to a sampling time T , by later stating that: "The minimum radionuclide concentration that can be measured with an accumulative sampler may be obtained from the following formula:

$$C = 3.7 \times 10^{10} q/TF,^* \quad (\text{Eq 76})$$

where C = detectable concentration, Bq/m³
 q = minimum detectable quantity of radionuclide, Bq
 T = sampling period, s
 F = sample flow rate, m³/s

One can attain any degree of sensitivity by increasing sample volume, TF , indefinitely" (excepting short-lived nuclides) (Reference 79, page 12).

An examination of the units shows that the above equation is in error and should simply be: $C = q/TF$. When corrected this equation may then be used to derive an acceptable minimum detectable amount (AMDA) of detectable activity, q_A , with an implied accuracy (precision) of 50% standard deviation at the AMDA level. For example, we first calculate the sampling rate needed to provide for an alarm when the total intake is 0.05 of an ALI (41). For C , we substitute $f \times \text{DAC}$ where DAC is the derived air concentration for a specific nuclide as given in Reference 41; DAC was originally called the Maximum Permissible Concentration (MPC) by the ICRP (40), and f is the fraction (or multiple) of DAC present as an air concentration for the specific nuclide.

The time $T_{0.05\text{ALI}}$ in seconds for a person breathing at a volumetric rate of 9,600 liters of air during an 8-hour work period (the breathing rate of Reference Man for light work activity (80)), in a concentration $f \times \text{DAC}$, to inhale an amount of a radionuclide equal to 0.05 ALI would be $0.05 \text{ ALI} / (f \times \text{DAC} \times 0.0096 \text{ m}^3/8 \times 3600 \text{ sec}) = 150,000 \text{ ALI} / f \text{DAC}$. For a breathing zone sampler (79) exposed to the same concentration as the person to give an alarm, or a positive indication of exposure with high probability, at the moment inhalation reaches 0.05 ALI. it would be sampling for the same sampling time T_s as the person; i.e., $T_s = 150,000 \text{ ALI} / f \text{DAC}$. If E is the fractional efficiency of collection of the sampler (e.g., filter paper), and F_s is the volumetric flow rate of the sampler, then the quantity of radioactive material collected during the period T_s , while the person is inhaling 0.05 ALI, is:

$$\begin{aligned} Q_s &= f \text{ DAC} \cdot T_s \cdot F_s \cdot E \\ &= 150,000 \text{ ALI} \cdot F_s \cdot E \end{aligned} \quad (\text{Eq 77})$$

*Editorial Note: Caution, this is an incorrect formula quoted from the original reference.

In order that the sampler detect the 0.05 ALI intake with a probability no greater than 0.05 of missing it (type II error), the following relationship must be maintained:

$$Q_s \leq q_A, \quad (\text{Eq 78})$$

where q_A is the MDA for the specific filter paper analysis procedures. For the amount of intake equal to 0.05 ALI to be measured accurately at a "determination level" (7) assuming perfect breathing zone sampling, the following must hold approximately true (7)

$$Q_s \leq 3q_A \quad (\text{Eq 79})$$

Thus, the required analytical MDA for detection of an 0.05 ALI intake (at P (Type II error) = 0.05) is $150,000 \text{ ALI} \cdot F_s \cdot E$ and the required MDA for a reasonably accurate determination of an intake of 0.05 ALI is at most $50,000 \text{ ALI} \cdot F_s \cdot E$. It is noted that since the sampler is operating for the same exposure time as the person, the MDA does not depend on the actual air concentration, but becomes more restrictive as ALI decreases and becomes more easily attainable as the volumetric flow rate F_s is increased. Also, by this analysis, the required MDA is found to be independent of sampling time, for a given breathing rate.

The "sensitivity" definition of Reference 79 ("...the lowest concentration that can be measured with a precision of $\pm 50\%$...") requires interpretation to be made consistent with Currie's definition (7) of limit of detection L_D (convertible to MDA by a calibration constant). For a decision level having a probability 0.05 of a Type I error, the L_D level of 4.65 s by Currie's derivation (7) would (for enough total counts so that Normal distribution statistics are valid) require a relative standard error of measurement of about $1/4.65 = 21.5\%$ at the L_D level.

If the "precision of $\pm 50\%$ " suggested in Reference 79 is taken to represent a 95% confidence interval about an estimated mean of $\pm 2s$, then the relative standard error in the net count of the activity would be about $50/1.96 = 25.5\%$. In order to interpret the IAEA sensitivity definition in a manner consistent with the usual definition of L_D (and MDA), we must then assume that for a measurement having $\pm 50\%$ accuracy at the 1.96s level, the L_D using Currie's symbol (7)) would need to at least satisfy the relation, $1.96 \cdot s/L_D = 0.5$ and $L_D = 3.92s$ (Currie (7) actually obtains 4.65s when considering in his derivation the Poisson variation of the standard deviation in blank counts, for a large blank count.)

If the "precision of $\pm 50\%$ " is interpreted to mean s/mean is about 0.5 when the detection limit L_D is reached, then an even less conservative L_D than Currie's (7) is obtained. In this case, we are led to assume that the IAEA standard intends to place $L_D = 2s$, which would give a Type II error probability of 0.36 at the 1.645s decision level (1) for a probability $\alpha = 0.05$ of a Type I error; a p (Type II error) = 0.36 is about 7 times higher than that used by Currie in his L_D definition (7).

From this example, it is clear that statements regarding accuracy in the national and international radiation safety standards and guides must be subject to considerable interpretation before definitions of accuracy, bias, precision, L_D and MDA can be obtained that are useful for specific ongoing quality control of radiobioassay programs.

Example: Detection Limit for Adequate Warning Time of Air Monitor

Suppose that an employee is suddenly exposed to the release of Pu-239 of Class W material that fills the laboratory with a continuous concentration 1000 times the DAC of 9×10^{-2} Bq/m³ (41). Assume there is no room air turnover to simplify the computation. The time T it would take a person to inhale 0.05 ALI (ALI = 200 Bq) when breathing at the Reference man rate of 9.6 m³/8-hours would be obtained from the equation

$$\frac{9.6 \text{ m}^3}{8 \text{ hrs}} \times T \times 9 \times 10^{-2} \text{ Bq/m}^3 \times 1000 = 0.05 \times 200,$$

which yields

$$T = 0.0925 \text{ hours} = 5.6 \text{ minutes} \quad (\text{Eq 80})$$

Assuming that a sampler of efficiency E and air flow rate F_s is required to alarm in the same sampling time $T_s = T$, then during this period of time, a sampler in the breathing zone would collect:

$$Q_s = F_s E T_s C \quad (\text{Eq 81})$$

$$\begin{aligned} \text{Assuming a filter paper with about 100\% efficiency and} \\ F_s = 400 \text{ liters/min} = 0.4 \text{ m}^3/\text{min} \text{ for a high volume sampler (81),} \\ Q_s^S = (0.4 \text{ m}^3/\text{min}) \times 1 \times (5.6 \text{ min}) (1000 \times 9 \times 10^{-2} \text{ Bq/m}^3) \\ Q_s^S = 201.6 \text{ Bq (about an ALI)} \\ Q_s^S = 12,096 \text{ d/min.} \end{aligned} \quad (\text{Eq 82})$$

Alpha counter detection limits are obviously not stringent for early detection of 0.05 ALI with high volume air sampling. A breathing zone or personal air sampler having a volumetric rate of about 1 liter/min (81) would collect $(1/400) \times 12,096 \text{ d/min} = 30 \text{ d/min}$ of ²³⁹Pu in the 5.6 min that it would take the wearer to inhale 0.05 ALI.

For an example of the estimation of an analytical detection limit (MDA), required to detect this small amount of ²³⁹Pu, assume that the alpha activity of the average "blank" sample obtained by running air samples in the laboratory before the plutonium release with use of the same measurement system and same waiting time for decay of radon daughters if required produced about 6 counts per minute. Also, assume that the standard counting time is 2 min to minimize the delay of detection of this (hypothetical) alarm-sampler. Then, the average blank count would be 12 counts, the standard deviation would be $\sqrt{12} = 3.46$ counts and the coefficient of variation of the blank count-rate would be $3.46/12 \times 100 = 29\%$. If another unknown sample were taken and the activity estimated by subtracting the activity of a paired measured blank, then the standard error of the estimated activity would be:

$$\sqrt{\text{Var}(C_U - C_B)} = \sqrt{C_U + C_B} \quad (\text{Eq 83})$$

If the unknown actually had no Pu activity present, the expected value of the sample standard error would be $\sqrt{2} \sqrt{C_B}$, and the coefficient of variation of a determination of net activity close to zero (due to Poisson counting error above) would be $1.414 \times 29\% = 41\%$. Fluctuations in the chemical yield of separation of the alpha activity from the filter paper may give a larger overall percentage error. On the other hand, if a standard constant average value for the blank is subtracted from the sample count to get the result, and chemical yield and calibration factors are relatively constant, then the variations of measurement may well consist mainly of the 29% standard deviation due to natural Poisson count fluctuations.

If the standard deviation of measurements due to all measurement fluctuations were, for example, double that due to the ± 3.46 count Poisson variation under the above assumptions, equivalent to ± 7 counts for a mean count of 12 in time T, and if the chemical yield (82) in recovering the plutonium for counting is 0.6, and the efficiency of detecting alpha particles in the detector per Pu-239 disintegration is 0.46, then the MDA value, q_A , would be about (see Section 2).

$$\begin{aligned} q_A &= 4.65 \times 7(0.60 \times 0.46 T) + 3/(0.60 \times 0.46 T) \\ &= 4.65 \times 7(0.60 \times 0.4 \times 2) + 3/(0.60 \times 0.46 \times 2) \\ &= 58.97 + 5.4 \\ &= 64 \text{ d/min, for a 2 minute count} \end{aligned} \quad (\text{Eq 84})$$

Unless counting time was lengthened, this would not quite be an adequate MDA to detect the 30 d/min activity collected in a breathing zone or personal air sampler in the 5.6 minute time to inhale 0.05 DAC at 1000 DAC ^{239}Pu concentration; however, it would be quite adequate to detect the activity on the corresponding high volume sampler for the same time period.

Actually, modern alpha spectrometers have much lower backgrounds, e.g., 0 to 4 counts in 2460 minute (83). Assuming a background (blank) count of 4 counts/2460 min, and (assuming that all background measurement fluctuations are zero) the first term in Eq (84) would become (taking ratios of standard errors)

$$\begin{aligned} q_{A1} & \text{ (alpha spectrometer)} \\ &= \frac{\pm \sqrt{4} \text{ counts}/2460 \text{ minutes}}{\pm 3.5 \text{ counts}/2 \text{ minutes}} \times .59 \\ &= 0.00046 \text{ d/min for a 2 minute count} \end{aligned} \quad (\text{Eq 85})$$

on an alpha spectrometer.

However, as shown in Section 2, for an essentially zero background or blank, if we want consistency in making p (Type II error) ≤ 0.05 , then the second term in Eq (84) would be dominant, and we would obtain

$$\begin{aligned} q_A &= 3/0.60 \times 0.46 \times 2 \\ &= 5.4 \text{ d/m as the lowest MDA for Pu-239} \\ & \text{for a 2-minute count on an alpha spectrometer.} \end{aligned} \quad (\text{Eq. 86})$$

Since the filter for the 1 liter/min sampler collected 30 d/min, these calculations show that an intake of 0.05 ALI is easily detectable on the 1 liter/min sampler if low background (alpha spectroscopy) is used and could warn the worker that he was being exposed at a rate of 1000 DAC and would receive 0.05 ALI if he continued to be exposed at the same rate for about another 90 minutes. Within this time period, rapid radiochemical procedures (84, 85) could possibly be used, to provide a check on the radiochemical species causing the alarm.

The above example was presented to indicate the relationships between lower limit of detection, L_D as defined by Currie (7), minimum detectable amounts (MDA) as defined in this standard, and the precision of measurements - as influenced by both counting and other procedural variations (20). Similar statistical considerations also prevail for the measurement of concentrations of radioactive material in biological samples.

Also, the discussion of the above example points out the many random variables that would be involved in air sample measurements, even if a perfect breathing zone sample could be obtained. However, even breathing zone sampling can often differ from the real exposure of man by a factor of 100 or more (Reference 79, p. 80). Thus, after-the-fact assessment of internal exposure of any significance must rely on bioassay and/or whole body counting (Reference 79, p. 4), even though air sampling is essential as a part of the monitoring program for early warning and assessment in facilities where large intakes are potentially possible in the event of otherwise undetectable releases of radioactive material. This relative importance of bioassay as the alternate means of individual dose assessment means that the accuracy requirements for bioassay measurements may need to be more stringent than those for air monitoring.

3.6.3.4 Legal Requirements for Accuracy

There are no explicit values given in laws or regulations for acceptable accuracy in the measurement or estimation of radiation doses to workers or general members of the public (86, 87). However, the need for some degree of accuracy is implied by the stated requirement that radiation exposure be monitored when "... it is likely that ..." persons will exceed 25% of the maximum permissible dose limits. Thus, greater accuracy of measurement is apparently necessary when persons approach exposures of 25% of permissible limits (e.g., 1.25 rem external radiation for a person whose annual limit would be 5 rems). Thus, one might infer that at a dose level of 1 rem, for example, it would be desirable that the upper 95% confidence interval boundary of an estimated exposure be no more than + 25% higher than the exposure estimate itself. This would provide an auditor with reasonable assurance that the 1.25 rem will not be exceeded, and that personnel monitoring would not be required if it could be assured that operations would remain unchanged and the exposure situation would remain constant. Of course, in practice, such situations would be rare and such assurances difficult to obtain.

An error limit of 25% at one rem would seem more stringent than the error limit of ± 1 rem recommended by the ICRP (55) (which would have a confidence interval ranging from 0 to 2 rem at the 1 rem level), or the factor of 2 uncertainty recommended by the NCRP (56) (which would correspond to a range of 0.5 to 2 rem, or -50% to +100%). Thus, a consistent rational approach to establishing accuracy limits can not be derived logically from present regulatory requirements and still be in agreement with some of the judgments of ICRP and NCRP committees.

Furthermore, the fact that the present and proposed regulations 10 CFR Part 20 incorporate maximum permissible concentrations (MPC's) (or "derived air concentrations") that are rounded off to one digit may be taken to imply that accuracy requirements on capabilities for regulatory purposes do not need to be more stringent than 33% (1.49 rounded to 1 gives about - 33% error).

In a similar way, no exact guidance on accuracy requirements of dosimetry may be derived from litigation history (88). Webster (89) has pointed out that accuracy requirements depend on the level of exposure to personnel, and to patients, in diagnostic radiology. He states, "The desirable accuracy is strongly influenced by the biological significance attributed to the dose and this is usually assumed to be dose dependent." He also points out that if the probability of induction of a harmful radiation effect is only 1% of the overall induction probability, then a factor of 10 of uncertainty in the dose level will not be important. If, however, the probability of radiation induction is 10%, then a factor of 10 uncertainty should be reduced if at all possible. Webster compares an estimated induction of leukemia per rad of whole-body dose of 2 per million per year with the normal incidence of about 65 per million per year in the U.S.A. to conclude that "... after receiving 1 rad of whole-body radiation there appears to be 1 chance in 30 that subsequent leukemia is radiation induced." Thus, if the whole body dose were 0.1 rad, an overestimate by a factor of 10 would not materially affect a conclusion that the illness was not radiation induced. However, if the dose were 10 rad and the error factor was 1/10 and gave an erroneous dose estimate of 1 rad, the correction of such an estimate would change the probability from "... a 1 in 30 chance to a 1 in 3 chance that the disease was radiation-induced." Since mean doses on the order of 10 rads are possible from diagnostic radiology, Webster concludes that "... an accuracy of dose estimation within a factor of 2 or 3 is desirable for all examinations which may contribute bone-marrow doses of the order of 1 rad." He reaches similar conclusions regarding doses to the fetus and newborn.

For personnel exposures in diagnostic radiology, Webster (89) concludes that the biologically significant doses are generally 10 or more times lower than the "skin" dose recorded by personnel monitoring devices, due to variations in photon energy with position relative to diagnostic tubes and shielding, the limited area of the beam usually intersecting the body, and the fact that the average marrow dose in the beam would usually be only about 10% of the dose to soft tissue near the skin surface. Some of these same considerations would make the biologically significant doses in radiation fields other than diagnostic radiology be less than those recorded for radiation protection purposes by a single personnel monitoring device. Of course, for radiation protection purposes, it is often deemed desirable to ensure that any errors in dosimetry tend to produce overestimates, rather than underestimates, of the effective whole-body doses. Thus, Webster's arguments would tend to indicate that personnel monitoring data would have limited value in litigation.

In recent years, an approach has been developed to "proving" for litigation purposes that a specific radiation exposure may be considered the cause of a disease or health event. This approach is sometimes referred to as the "probability of causation" approach (90). Although this approach has still not received general acceptance in the courts as a valid formulation of the probability of causation (91), it has been considered valid in one recent court decision (92), and may be examined for its implications regarding criteria for

accuracy in dosimetry. The approach is analogous to Bayes methods of determining a posteriori probabilities in terms of a priori probabilities (93), and may be examined by considering the following equation:

$$\Pr(E_j/F) = \frac{\Pr(F/E_j) \Pr(E_j)}{\sum_{i=1}^n \Pr(F/E_i) \Pr(E_i)}, \quad (\text{Eq 87})$$

where the slant (/) here represents the conditional line, meaning "given the event..", and not the division sign. This equation is presented as Equation (6.10) of Reference 94, where the symbols are defined in more detail, and two examples of its application are given. Here, the symbols may be briefly defined as

- $\Pr(E_j/F)$ = the probability that event E_j will occur given that F has occurred,
- $\Pr(F/E_j)$ = the probability that F will occur given that E_j has occurred,
- $\Pr(E_j)$ = the probability that E_j will occur.

The events E_i are assumed to be mutually exhaustive as well as mutually independent in the derivation of the above equation. By examination of the second example of the use of this equation in Reference 94, an analogous relationship may be set up as follows:

$$P(\text{radiation was the cause, given cancer occurred}) = A / A + B, \text{ where} \quad (\text{Eq 88})$$

$A = \Pr(\text{cancer occurs/radiation exposure}) \Pr(\text{radiation exposure})$
 $B = \Pr(\text{cancer occurs/all other causes}) \Pr(\text{all other causes occurred})$ and

it is presumed that a specific amount of radiation exposure occurred. This relationship obviously oversimplifies the actual situation where various levels of radiation exposure may occur, and the events of exposure to the different causative agents of a disease may not be independent or mutually exclusive. Not only may a person be exposed to more than one agent at a time, but also the known interactions of agents such as co-carcinogens may influence probabilities of disease occurrence in complex ways depending on the exact nature and quantities of the agents.

Although Bond (90) acknowledges that his approach need not be referenced to the Bayesian method, and that there may be uncertainties in knowledge of exposure to the other agents, he suggests an approach equivalent to the above relationship as a first approach to objective judgments in radiation litigation cases. In effect, the lumping of all other agents into a single term in the denominator would effectively bias the calculated probabilities that specific measured radiation exposures have caused specific cases of disease in the direction of overestimating these probabilities. This would be true if for no other reason than that the other causative agents are not likely to be identified and measured, so that an average population risk replaces the second term in the denominator while attention is focussed on the particular measured radiation exposure of the individual under consideration.

Nevertheless, the above approach seems to be under broad consideration for use in court cases and has received a limited acceptance, so it is of interest to examine the implications on accuracy. By examining Bond's example of page 109(90) where he obtains a relative attributable risk of 28% for an exposure of 10 rad (when the normal leukemia incidence is assumed to be $25 \times 10^{-6}/\text{yr}$), then it is evident for this example that a dose of 25 rad would be judged to give a relative attributable risk of 50%. Thus, any slightly greater dose than 25 rad could by this method be judged in court to have "... more probably than not" been the cause of a particular case of leukemia. Thus, if this method should prevail, then it would be extremely important to be able to prove, e.g., that a recorded dose of 20 rad had an error of about 2 standard deviations that was less than 5 rad, or $\sigma \leq 2.5$ rad at the 20 rad level. Accuracy requirements according to this simple criterion would then be dependent upon dose level and exact, specified assumptions regarding the normal incidence of given diseases versus age and other parameters, and exact assumed values for radiation risk coefficients. Thus, again, this method probably has little likelihood of offering guidance for the establishment of a generic specification for the bias and precision of dose or bioassay measurements.

Another approach to examining the question of radiation causation appeared convincing to a jury in the case of Dennis versus the Department of Energy/General Electric Company (95). In this case, the concept of doubling dose was explained verbally as well as mathematically in simple formulae, and it was acknowledged that other agents of potential causation were known to have been present, but not measured nor of exactly known etiology. Thus, the resulting argument reduced to that of showing that the dose to the individual from the job was 2.2 rem compared to the equivalent of 11 rem of background radiation (0.2 rem/year, including the contribution of 0.1 rem/year from the radon and daughters' weighted dose equivalent exposure as presented in the most recent United Nations report after conversion to conventional units (96). Thus, according to this approach, the probability must have been less than $2.2/(2.2 + 11) = 16\%$ that the disease in question was caused by the occupational exposure. Since the jury seemed to accept these arguments by virtue of its decision, it appears that it may be important in radiation dose determinations to ensure that the uncertainty of measurement of the dose in a given year should be small compared to the difference between total exposure and background radiation exposure. This again provides a variable benchmark for required accuracy. As a result of experience with this case as well as others, the defending attorney has suggested that it may be of value in litigation to have determined exposures below the 5 rem per year level to within 0.1 rem, but that it is probably not necessary to achieve accuracies of 0.01 rem; this opinion is consistent with the needs of litigation (91). The fact that the dose in the Dennis case had been measured by a number of badges on a belt, and that the badges had agreed within 0.1 to 0.2 rem appeared to impress the jury, and this seems to support the attorney's opinion.

The use of Relative Attributable Risk (RAR) for litigation cases has been shown to be equivalent to the use of a relative risk assessment, and in both cases highly variable risk estimates would lead to considerable uncertainty in the calculated result (97). Katz (4) has discussed the view that uncertainties, in the form of confidence interval estimates, are presentable to a judge, jury and counsel, but stresses that the expert witness will be required to explain his data collection procedures in detail (in our case, those would include measurement methods).

Despite the recognition that exposure information is often imprecise and risk estimates are uncertain, Tinsley (98), after a review of 250 Workmen's Compensation cases based on radiation exposure still concludes as follows: "I need not tell you that the better the basic information; the more adequate, accurate and reliable the exposure data; the better the conclusions and the decisions will be in a particular case. Conscientious, careful, and professional measurement of the quantum of exposure received both medically and industrially protects the interests of the employee both as an employee and human being and serves equally well to protect the interests of an employer or other interested party. The only point remaining is to decide if the time, trouble and cost involved is worth it in terms of the results we might hope to achieve." Also, O'Toole (99) states that, "... I wish only to indicate that in the courtroom evidence which is presented in mathematical terms tends to receive a degree of respect in excess of what the underlying facts would warrant." One of Huard's (100) conclusions is: "The scientific uncertainty of radiation exposure measurement will open the record of such measurement to vigorous, and often successful, challenge in court." Kronzer (101) reports a decision of the Texas Supreme Court in January 1969 which would make it "... probable that our court would hold that a case for leukemia or cancer could not be made on the basis of probabilities of less than 51 percent." These and other papers presented at the 1969 conference (88) indicate that while the accuracy requirements of radiation dosimetry for purposes of litigation are uncertain and not likely to be readily standardized, more often than not the experienced attorneys seem to feel that more accurate and reliable dosimetry is likely to be a positive influence in the administration of justice. Similar conclusions can be drawn from more recent litigations (102-104), which also indicate that it may be particularly important for defendants responsible for dosimetry to be able to show proper quality control of measurements and regular participation in inter-laboratory tests and accreditation programs.

The more quantitative specification of accuracy needs of dosimetry for litigation and regulatory purposes will probably be possible only after Congress resolves the approach (105) to deciding just restitution for persons inflicted with diseases for which only probabilistic descriptions can be given in regard to their causation by environmental agents. Indeed, it seems that our Society needs to reduce its uncertainty regarding the definition of "justice" in such cases before we can reduce our uncertainties in determining accuracy requirements of dosimetry for litigation purposes.

3.6.3.5 Accuracy Requirements in Management of Emergencies Involving Radionuclide Intake.

Although there seems to be no general guidance on the quantitative accuracy needs of bioassay for emergency use, and it is clear that errors of a factor of three or more in estimating intakes may have to be accepted at early times, if not at later times, after accidental intake (106, 107), accident experience does offer some considerations that suggest limitations on errors of emergency bioassay measurements. Early estimation of intakes from bioassay measurements are often requested by the physician managing the case both to: (1) determine if the likely internal exposures are high enough to warrant chelation or other treatment, and (2) if trial chelations are initiated, to determine the efficacies of these treatments (68, 69). In the first case, it is important to obtain immediate results, usually accurate to within a factor of 2; but the requirement for rapid turnover requires special bioassay procedures that may correspond to

higher than usual MDA's (84, 85). During chelation therapy for removal of plutonium, americium or other heavy elements (the more common nuclides to require chelation), the immediate requirement is to obtain an estimate of efficacy before the next administration is required. This could require analysis results within less than several days for weekly administrations over a long time period (108), or even more frequent analyses in early times after intake during more intensive chelation treatment (69, 109). Thus, rapid methods are also needed for evaluation of chelation efficacy, i.e., the ratio between excretion rate the day after treatment to the excretion rate that would have been expected without treatment.

The efficacy of different chelates may differ, sometimes at different phases of the therapy, but for plutonium, an efficacy of 50-100 is often assumed for DTPA administration (110, 111). Chelation treatments of cases of ^{241}Am inhalation exhibit efficacy averaging about 5 to 8 over the first three years after intake (112, 113) and an estimated efficacy of about 10 over a 7-year period (108). After some individual 1-gram doses of DTPA, the increased excretion of Am-241 was only about a factor of 3 above the excretion the day before treatment (113). A 30-year old male of about standard-man size who provided total daily urine samples for a period of many months exhibited a daily volumetric output that was approximately log-normally distributed with mean 840 ml and a standard geometric deviation of $s_g = \frac{x}{\bar{x}} = 1.4$ (112).

Thus, an implied requirement for accuracy of urinary bioassay can be derived from the need to detect increased daily outputs resulting from DTPA: it would be desirable to be able to detect a 3-fold increase in output with greater than 95% confidence. Now, the product of two log-normally distributed variates is itself lognormally distributed and the variance in the logarithmic quantity resulting is compounded as follows (Reference 51, p. 11):

$$\text{Var} (\ln x) = \text{Var} (\ln x_1) + \text{Var} (\ln x_2) \quad (\text{Eq 89})$$

which is the same as

$$(\ln s_g)^2 = (\ln s_{1g})^2 + (\ln s_{2g})^2, \quad (\text{Eq 90})$$

where s_g represents the corresponding "standard geometric deviation" (52). If we consider that an efficacy determination is the ratio of two lognormally distributed bioassay measurements, then the efficacy measurement is lognormally distributed since the inverse of a lognormally distributed variate is lognormally distributed with the same s_g (Reference 51, p.10). Further, assuming that each bioassay determination involves the division of a count or count-rate by a calibration constant (or recovery) and a volume of urine sampled, then we can estimate the required precision of analysis to meet the requirement that an efficacy of 3 be detected with high probability.

If we set a factor of 3 as the upper level of a lognormal 95% confidence interval (two-sided for convenience of calculation), then the value of s_g for the efficacy determination should be $3^{1/1.96} = 1.7516$. Then, using the above variance relationships, the s_g of a single bioassay determination should be held to:

$$\begin{aligned}
 (\ln 1.7516)^2 &= 2 (\ln s_g(\text{bioassay}))^2, \text{ or} \\
 \ln s_g(\text{bioassay}) &= \ln 1.7516 / (2)^{\frac{1}{2}} = 0.5605 / 1.414 = 0.3964. \\
 \text{and } s_g(\text{bioassay}) &= \exp(0.394) = 1.4865. \qquad \qquad \qquad (\text{Eq 91})
 \end{aligned}$$

This standard geometric deviation represents the following confidence interval about the median μ_g :

$$\begin{aligned}
 (\mu_g / 1.4865, 1.4865 \mu_g) &= (0.6727 \mu_g, 1.4865 \mu_g) \\
 \text{or } \underline{\mu_g - 33\%, + 50\%}, &\text{ approximately.} \qquad \qquad \qquad (\text{Eq 92})
 \end{aligned}$$

If the variability of urinary output is incorporated in the uncertainty of the analytical determination and not corrected for (e.g., by creatinine determinations), then the analytical procedure may need to limit the combined errors due to analytical variations in yield, counting statistics, etc. to well within -33%, 50%. (Also, the analytical techniques for use in efficacy determinations should be applicable to measurement of americium still complexed with DTPA after excretion (112, 114.) If the distribution of these other combined errors of analysis can be assumed to be approximately lognormal, then the above variance relationships can be used to estimate the required precision for the analytical method alone.

3.6.3.6 Accuracy Needs for Epidemiologic Investigations of Radiation Risk.

As pointed out in previous sections, radiation dose estimates from personnel monitoring devices and bioassay measurements have often been purposely biased on the high side when necessary to compensate for uncertainties of measurement or dose interpretation from practical field measuring devices or methods (115, 116). However, it is possible and desirable for purposes of epidemiologic investigation to remove considerable bias from personnel monitoring records by examining original methods of dosimetry in the respective installations over time, and converting exposures to common units by a common interpretation (116, 117, 118). Generally, dose estimates should be unbiased for epidemiologic use, and the ranges of uncertainty in individual and population dose estimates should be recorded for use in estimating the uncertainties, in the conclusions of such studies (116, 118). Random errors in individual readings of personnel monitoring devices, and even a limited proportion of lost or omitted dose values, can often be compensated-for in estimating doses to larger populations by proper evaluation and recording of data so that the law of large numbers can average out the random variations of measurement and interpretation (2, 115, 116, 66). Thus, in terms of accuracy needs for prospective monitoring programs, the need to reduce bias mistakes in measurement and recording, and lost data, may be of greater importance than the need to reduce random errors that have a central tendency about the "true" value of dose at the point(s) in tissue of interest.

The quantitative specification of accuracy needs for a particular epidemiologic investigation would be dependent on the precision to be expected in estimating differences in health effects between populations in a given study, if indeed any differences are detectable. For example, Jablon has estimated that overall changes of about a factor of 2 in dose estimates for the atomic bomb survivor

study would probably change the central estimates of radiation risk by about a factor of 2 (119). However, the 95% confidence limits on risk estimates, even for a given dose-response model, can vary by more than a factor of two at any given dose level (120).

In regard to the radiation worker population, the entire AEC Contractor radiation worker population might need to be followed for several generations to even detect with statistical confidence the more prevalent radiation-induced diseases (121). Even then, what might seem to be minor differences in matching employee to "control" populations could completely obliterate either the significance, or lack of significance, of small observed differences in mortality, which need to be properly determined and tested in order to assess risks of radiation-induced disease from occupational levels of exposure (122, 123, 124).

3.6.4 Conclusions Regarding Accuracy Requirements of Bioassay, Specifications of Bias and Precision Statistics and Acceptable Ranges for Sample Quality Control Tests

3.6.4.1 General Conclusions

The foregoing review indicates the ranges of accuracy that might be considered acceptable for various purposes and the need to develop de nouveau easily applied performance specifications of accuracy specific to the purposes of this standard. There is evidently a general value to improved accuracy of bioassay for many purposes; however, accuracy limits for performance testing of sample measurements that are too tight can be counterproductive to radiation protection purposes for several reasons:

1. Too great an expenditure of time and expense for individual sample measurements - beyond practical needs for accuracy - can result in an inadequate number of bioassay samples to monitor particular operations.
2. Unnecessary procedural complexities or unduly long counting times can result in delayed reporting of bioassay results, decreasing probabilities of timely correction of unwarranted or unacceptable exposure situations.
3. Improvements in single sample measurements beyond required precision limits may not be meaningful, since dose or dose commitment estimates from limited numbers of samples may be in error by factors of 4 just from uncertainties in retention functions and variations of factors of 2 in individual urinary excretion rates (46, 112). (Still, quality assurance programs should be designed to reduce bias (systematic) errors so that long term follow up of the more highly exposed individuals will result in improved estimates of internal radioactivity burdens and dose commitments (112, 113, 108). As in air sampling, the number of repeated measurements needed to "average out" random errors to a given degree of precision is dependent upon the dispersion of individual measurements (125).

The general conclusions regarding accuracy requirements may be summarized from the foregoing literature review and analysis as follows:

1. Accuracy needs of dosimetry for radiation protection programmatic purposes have been expressed in varied and only general form by recognized expert bodies, and are not directly applicable to specification of sample bias and

precision performance criteria. Generally, better accuracy is always desirable, but when dosimetry systems must incorporate error over certain ranges of radiation quality or energy, it is usually deemed better for radiation protection purposes that any uncertainties be biased on the high (+) side of the true dose.

2. Administrative and legal requirements would also make the reduction of all errors desirable, when cost effective. However, here again a bias on the high side is desirable when uncertainties of measurement or interpretations are present in order to be able to state with a high degree of confidence that regulations are indeed complied with, or that cumulative exposures and risks are below a certain given level.
3. Accuracy requirements for emergency dosimetry may be less stringent during early phases in order to allow use of more rapid bioassay methods, applicable to higher initial radioactivity levels or concentrations. However, when chelation treatments are to be considered, random errors must be limited to a standard geometric deviation of x or $\div 1.5$ or less (-33%, +50%) in order to allow detection of chelation efficacy, for the more radiotoxic nuclides to which workers have been exposed.
4. Accuracy needs for epidemiologic investigation vary depending on the expectations of accuracy in mortality or morbidity differentiation and the accuracy needs for radiation risk estimation in developing basic radiation protection standards.

In general, it is most important that methods of dosimetry and plant exposure conditions, as well as dosimetry results, be documented and retained in archival quality for several generations, so that unbiased estimates of individual and population sub-groups can be retrospectively obtained by corrections of the dose data that are conservatively estimated for radiation protection purposes.

3.6.4.2 Specific Conclusions Regarding the Selection of Specific Accuracy Criteria and Practical Statistical Estimators

In view of: the need to balance the above considerations, the fact that restricting average bias and precision estimated from small sample sizes to a finite range does indeed require more stringent accuracy for the underlying population of measurements (see the remainder of this section and Table 4); the greater detriment of negative than positive errors of bias for administrative, legal, and ALARA risk limiting purposes of radiation protection; and the fact that the bias and precision testing criteria selected by the ANSI N13.30 working group have now been tested (22) and shown to provide an appropriate balance between the need to bring deficient performance up to state-of-the-art standards and the need to avoid failing laboratories that have satisfactory overall performance, the following bias and precision criteria selected by the working group (24) are endorsed by this author for incorporation into the performance standard.

$$-0.25 \leq B_r \leq 0.5 \quad (\text{Eq 93})$$

$$|S_B| \leq 0.4 \quad \text{and} \quad |S_A| \leq 0.4 \quad (\text{Eq 94})$$

However, as is evident from the results of the performance testing study (22, 23) as well as the recent discussion by Currie (20), the above criteria are deemed applicable to performance testing and intralaboratory quality control only when samples are spiked with quantities at least 4 times the minimum detectable amount (MDA), calculated as described in Section 2. Of course, if the acceptable MDA (AMDA) for purposes of this standard (which includes needs and cost-effectiveness considerations for radiation protection purposes) is 4 MDA or more, it may be appropriate to use the above sample bias and precision criteria for testing at ≥ 1 AMDA. Separate test criteria are provided in the body of this standard for determining whether a laboratory is meeting the performance criteria that their MDA for a given respective measurement is an $MDA \geq AMDA$.

3.7 Implications of Finite Sample Sizes in Terms of Failing the Precision Criteria*

Let x be a random variable with mean μ and variance σ^2 , representing the measured activity value. Let "a" be the known activity added. Then, $b = \frac{x - a}{a}$ is a random variable with mean $\frac{\mu - a}{a}$ and variance σ^2/a^2 . Let

$$s^2 = \frac{\sum (b_i - \bar{b})^2}{n - 1} \quad (\text{Eq 95})$$

where n is the sample size and $\bar{b} = \sum b_i/n$.

Now, if x is distributed Normally, then so are the statistics b and \bar{b} . Then, the statistic $\frac{(n-1) s^2}{\sigma^2/a^2}$ is distributed as a chi-square variate with $(n-1)$ degrees of freedom.

If the precision criteria of the performance test is failed when $|s| > 0.4$, then we need to know what is the probability of observing $|s| > 0.4$ for a finite number of samples in a test category, when the population relative standard deviation is σ/a . (The small "s" as defined by Equation 95 is the same function as S_B , in the symbol used to express S_A/A_a --see Sections 3.4 and 3.5. The small "a" becomes A_a in the symbolism of Section 3.4.)

Let α = the probability of failure, given σ/a

$$y = \frac{(n - 1) (0.4)^2}{\sigma^2/a^2} \quad (\text{Eq 96})$$

$$m = n - 1$$

Then α is found as the area under the chi-square (χ^2) density function, above the value of y , and

*This approach was suggested by Dr. Charles T. Schmidt, University of California, consultant to the ANSI N13.30 Working Group, who also provided the analytical discussion of Equations 95-97 and Table 1.

$$\int_{z=0}^y \frac{1}{2 \Gamma(m/2)} (z/2)^{\frac{m}{2} - 1} e^{-z/2} dz = 1 - \alpha \quad (\text{Eq 97})$$

This may be solved explicitly for m even. The following table shows values of $\alpha = P(s > 0.4, \text{ given } \sigma/a)$ for some values of n and σ/a :

Table 1

Probabilities of Failing Precision Test Versus Relative Error and Sample Size

n	3	5	7
σ/a			
0.2	0.018	0.003	0.001
0.3	0.17	0.13	0.099
0.4	0.37	0.41	0.42
0.6	0.64	0.78	0.85
0.8	0.78	0.91	0.96

Thus, to pass the test, "usually", the population σ/a (i.e., the true underlying relative precision of the measurement) should be about 0.25 or less. Note also that the probability of passing when $\sigma/a = 0.8$ and $n = 3$ is $(1 - 0.78) = 0.22$.

Tables of integrals (126) equivalent to the cumulative probability under the chi-square function of Equation (97) may be used to calculate the required relative precision of the underlying measurement in order for all samples in a category with N measurements to pass the relative precision criteria $|S_B| \leq 0.4$ of Section 3.6.4.2. Table 2 below presents the true relative precisions required in order to have 95% and 99% chance of passing the criteria $|S_B| \leq 0.4$ for various sample sizes.

Table 2 indicates, when the sample size is $N = 3$ for example, that for a laboratory to pass the precision criteria of $|S_B| \leq 0.4$ for the calculated relative precision statistic of Section 3.6.4.2, its true precision must be 23.1% to pass 95% of the time. Since $S_B = s_A/A_a$, the expression $(N-1)S_B^2/(\sigma^2/A_a^2)$ becomes $(N-1)s_A^2/\sigma^2$, by cancelling A_a . In the same way, by cancelling \bar{A} , it can be shown that Table 2 can be used to determine probabilities of $|S_A| \leq 0.4$ when σ/A_a (or σ/a) is replaced by s_A/\bar{A} , the coefficient of variation.

3.8 A Computer Program for Calculating the Probability of Passing the Precision Criteria for a Replicate Test Measurement of N Spiked Samples

A computer program written for the HP 41CV, but adaptable as an algorithm for other micro-computers, is presented in Exhibit 1 for use in calculating for any replicate of N measurements, each spiked with an amount $A_a = AA$, the probability of passing the precision test, $|S_B| \leq 0.4$ for a known distribution of underlying population values having standard deviation $\sigma = \text{SIG}$. The program allows for

Table 2

True Analytical Precision Necessary to Pass Relative Precision Criteria $|S_B| \leq 0.4$,
For Various Sample Sizes

Sample Size Number of samples in given test category,	Degrees of Freedom,	$\chi^2_{0.95}$ Value for 95% Chance of passing*	True Relative Precision limit for 95% chance of passing,	$\chi^2_{0.99}$ Value for 99% chance of passing*	True Relative Precision limit for 99% chance of passing,
N	n = N-1		$\sigma/a = \frac{((N-1)(.4)^2)^{1/2}}{\chi^2_{0.95}}$		$\sigma/a = \frac{((N-1)(.4)^2)^{1/2}}{\chi^2_{0.99}}$
2	1	3.84	0.204 (20.4%)	6.63	0.1553 (15.5%)
3	2	5.99	0.231	9.21	0.186
4	3	7.81	0.248	11.3	0.206
5	4	9.49	0.260	13.3	0.219
6	5	11.1	0.269	15.1	0.230
7	6	12.6	0.276	16.8	0.239
8	7	14.1	0.282	18.5	0.246
9	8	15.5	0.287	20.1	0.252
10	9	16.9	0.292	21.7	0.258
15	14	23.7	0.307	29.1	0.277
20	19	30.1	0.318	36.2	0.290

*Applicable also to criteria $|S_A| \leq 0.4$ when σ/A_a is replaced by σ/\bar{A} .

substituting other precision criteria (SL = precision limit) in the event a different criteria for the acceptable range of the precision statistics is used. Also included in Exhibit 1 is a subroutine for calculating the value of the Gamma Function for DF/2, where DF is the number of degrees of freedom N - 1. This Gamma Function is needed to calculate the integral under the Chi-Square probability density function for DF degrees of freedom, which when integrated up to the limit y of Equation (96), $y = \frac{(N-1)(0.4)^2}{SIG^2/AA^2}$, in symbols adapted for the computer program, will give the probability of the replicate statistic S_B passing the precision test. Programs are written to allow flexibility in selection of parameters for use in testing the programs over various parameter ranges, as well as to yield the probability of passing the precision test.

Tests of the program's calculations are presented in Table 3 for comparison with the probabilities in Table 1. The programs, when carried out for abscissa interval lengths of 0.05, are seen to provide answers in Table 3 that are usually within 1% of those in Table 1. The computer program results have been rounded off to four digits for possible comparison with future test calculations. The Gamma Function subroutine is actually more accurate than 1% when compared over the range of values available in mathematics and statistics handbooks.

EXHIBIT 1 - Program for Calculating Chance of Passing Precision Tests and Subroutine for the Gamma Function

FFF *FCHISQ*	54 *N=18L2/SIG2/AR*	107 STO 09	158 PSE	
	55 APCL Y	108 0	159 *PSE*	
01*LBL *FCHISQ*	56 AVIEW	109 STO 13	160 PSE	
02 *THIS PROGRAM*	57 PSE		161 BEEP	
03 AVIEW	58 PSE	110*LBL 03	162 ENT	
04 *CALCULATES FCHISQ*	59 *PSE*	111 RCL 09		
05 AVIEW	60 *ENTER UL OF*	112 2		
06 *SQUARE OF P OF*	61 AVIEW	113 /		
07 AVIEW	62 *INTEGRATION*	114 CHS		
08 *SP LESS THAN*	63 AVIEW	115 E*Y		
09 AVIEW	64 PSE	116 STO 10		
10 *PRECISION LIMIT*	65 *0 =N-18L2/SIG2*	117 RCL 16		
11 AVIEW	66 AVIEW	118 2		
12 *P SA LT SL*	67 *AP, OF OTHER *	119 -		
13 AVIEW	68 AVIEW	120 2		
14 *USING GAMMAFN *	69 *CHOICE = 1*	121 /		
15 AVIEW	70 AVIEW	122 ENTER*		
16 *SUBROUTINE*	71 PSE	123 RCL 09		
17 AVIEW	72 *CHISQUL=0.1?*	124 X<Y		
18 *N = ?*	73 PROMPT	125 Y*Y		
19 PROMPT	74 X=0?	126 STO 11		
20 STO 01	75 GTO 01	127 RCL 16		
21 *STD DEV OF MEAS*	76 GTO 02	128 2		
22 *UPPERMNT SIG*		129 /		
23 *SIG = ?*	77*LBL 02	130 ENTER*		
24 PROMPT	78 *CHOOSE UPPER *	131 2		
25 STO 02	79 AVIEW	132 X<Y		
26 *ENTER REL PRECI*	80 *LIMIT*	133 Y*Y		
27 *SIGN LIMIT*	81 AVIEW	134 STO 12		
28 *SL =.4 OF ?*	82 PSE	135 RCL 09		
29 PROMPT	83 *PSE*	136 RCL 10		
30 STO 03	84 *CHISQUL=0*	137 *		
31 *TRUE VALUE*	85 *PROMPT*	138 RCL 11		
32 *AR = ?*	86 STO 15	139 *		
33 PROMPT		140 RCL 07		
34 STO 04	87*LBL 01	141 /		
35 *CALCULATE UPPER*	88 RCL 01	142 RCL 12		
36 *LIMIT OF INTEGR*	89 1	143 /		
37 *ATION,CHISQUL*	90 -	144 ST+ 13		
38 RCL 01	91 STO 16	145 RCL 08		
39 1	92 *DF =*	146 ST+ 09		
40 -	93 ARCL X	147 RCL 09		
41 RCL 02	94 AVIEW	148 ENTER*		
42 ENTER*	95 PSE	149 RCL 15		
43 *	96 PSE	150 X<Y?		
44 /	97 PSE	151 GTO 04		
45 RCL 04	98 XEQ *GAMMAFN*	152 GTO 03		
46 ENTER*	99 STO 07			
47 *	100 BEEP	153*LBL 04		
48 *	101 *DELX=.01 OF .05*	154 RCL 17		
49 RCL 07	102 PROMPT	155 *FCHISQ =*		
50 ENTER*	103 STO 08	156 APCL X		
51 *	104 RCL 08	157 AVIEW		
52 *	105 2			
53 STO 15	106 /			
			FRP *GAMMAFN*	41 Y*Y
			01*LBL *GAMMAFN*	42 STO 07
			02 0	43 RCL 06
			03 STO 01	44 *
			04 0	45 RCL 04
			05 STO 11	46 *
			06 0	47 STO 08
			07 STO 06	48 ST+ 09
			08 .000001	49 RCL 04
			09 STO 12	50 ST+ 05
			10 FIX 2	51 RCL 09
			11 *DF=?*	52 RCL 10
			12 PROMPT	53 -
			13 STO 02	54 STO 11
			14 RCL 02	55 ENTER*
			15 ENTER*	56 RCL 12
			16 2	57 X<Y?
			17 /	58 GTO 02
			18 STO 03	59 GTO 01
			19 *D=.01 OF .05?*	60*LBL 02
			20 PROMPT	61 RCL 09
			21 STO 04	62 ENTER*
			22 RCL 01	63 .8
			23 ENTER*	64 X<Y?
			24 RCL 04	65 GTO 07
			25 2	66 GTO 01
			26 /	
			27 +	67*LBL 03
			28 STO 05	68 RCL 07
			29*LBL 01	69 *DF/2=*
			30 RCL 09	70 ARCL X
			31 STO 10	71 AVIEW
			32 RCL 05	72 PSE
			33 CHS	73 *PSE*
			34 E*Y	74 RCL 09
			35 STO 06	75 *GAMDF/2=*
			36 RCL 05	76 APCL X
			37 ENTER*	77 AVIEW
			38 RCL 03	78 PSE
			39 1	79 *PSE*
			40 -	80 PSE
				81 ENT

Table 3

Calculations of the Values of $1 - FCHISQ$, the Probability of Failing the the Precision Test, in Table 1, Using the Computer Programs of Exhibit 1

SIG/AA	N =		
	3	5	7
0.2	0.01822	0.003076	0.0005113
0.3	0.1694	0.1307	0.09980
0.4	0.3678	0.4060	0.4232
0.6	0.6376	0.7725	0.8513
0.8	0.7788	0.9098	0.9595

Again, by considering that values of SIG/AA represent values of the statistic $S_A = s_A/A$, the same computer programs can be used to calculate the probabilities that $|S_A| \leq 0.4$.

3.9 Implications of Finite Sample Sizes in Terms of Passing the Bias Criteria

3.9.1 Probabilities to be Calculated

The task is to calculate the probability that a bias estimator, B_r , for a finite sample of N measured test samples, lies within a specified range, taken as 0.25 to 0.50 in the draft ANSI N13.30 standard (24). The bias estimator is defined, for reasons given in Sections 3.1-3.2 as

$$B_r = \bar{B}_{ri} = \frac{1}{N} \sum_{i=1}^N B_{ri} \quad (\text{Eq 98})$$

$$= \frac{1}{N} \sum_{i=1}^N \left(\frac{A_{ai} - A_a}{A_a} \right)$$

Thus, in symbols, we wish to calculate $P(b_1 \leq B_r \leq b_2)$. The remainder of this Section will describe the determination of P for two cases: a) normally (Gaussian) distributed values of A_{ai} , the measurements, and b) log-normally distributed A_{ai} . One or the other of these two cases can usually be used in practice to represent the frequency distribution of measurement values, or at least an approximate envelope to the frequency distribution. The log-normal function may be expected to be more representative of measurement distributions when the measurement results are obtained or calculated by procedures that would exclude $A_{ai} < 0$, or when an approximate continuous distribution is to be fitted to discrete Poisson-distributed data with mean values of about 50 or more.

3.9.2. Probability Statements for Normally-Distributed Values of A_{ai}

For normally-distributed A_{ai} , the probability density function for A_{ai} is the well-known error function:

$$p(A_{ai}) = \frac{1}{\sqrt{2\pi} \sigma} e^{-\frac{(A_{ai} - \mu)^2}{2\sigma^2}} \quad (\text{Eq 99})$$

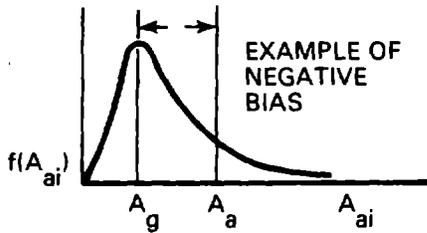
and
$$p(a_1 \leq A_{ai} \leq a_2) = \int_{a_1}^{a_2} p(A_{ai}) dA_{ai}, \quad (\text{Eq 100})$$

where A_{ai} is taken here to represent the (continuous) measured variable.

Since B_{ri} and B_r are linear functions of A_{ai} , then they are also normally distributed when the A_{ai} are normally distributed. However, B_{ri} and B_r are not

necessarily normally nor log-normally distributed when the A_{ai} are log-normally distributed. Thus, in order to develop parallel methods of determining $P(b_1 \leq B_r \leq b_2)$ for both cases, we will consider the normal distribution of A_{ai} (and $\log A_{ai}$ when A_{ai} is log-normally distributed) and its associated parameters μ (the mean) and σ (the standard deviation), and not the corresponding distributions of B_r and B_{ri} .

3.9.3 Probability Statements for Log-Normally Distributed Values of A_{ai}



Let the individual measurements of a laboratory be distributed log-normally, with probability density function $f(A_{ai})$,

such that
$$\int_0^{\infty} f(A_{ai}) dA_{ai} = 1, \quad (\text{Eq 101})$$

and
$$f(A_{ai}) = \frac{1}{\sqrt{2\pi} \sigma_g} e^{-\frac{(\ln A_{ai} - \ln \mu_g)^2}{2\sigma_g^2}}, \quad (\text{Eq 102})$$

where σ_g is the standard deviation in the frequency distribution of $\ln A_{ai}$ (the natural logarithms of the A_{ai} values) and $\ln \mu_g$ is the mean value of $\ln A_{ai}$, μ_g being defined as the "geometric mean" of A_{ai} (52). Now, let X represent the variable of the underlying measurement, of which A_{ai} are sample values, and let x represent a specific value of A_{ai} for ease in analysis.

Then, the cumulative probability distribution of $\ln x$ becomes:

$$P(\ln X \leq \ln x) = \int_{\ln X = \ln 0 = -\infty}^{\ln X = \ln x} f(X) d(\ln X) \quad (\text{Eq 103})$$

The cumulative probability distribution in X rather than $\ln X$ is obtained from the above equation by substituting for $d(\ln X) = dX/X$, and changing limits appropriately:

$$P(X \leq x) = \int_{X=0}^{X=x} \frac{f(X)}{X} dX \quad (\text{Eq 104})$$

The integrals are, of course, defined as the limits of the corresponding integrals over the domain (E, x) as $E \rightarrow 0$, for $0 < E < x$.

The underlying population mean of X is

$$\bar{x} = \int_{X=0}^{X=\infty} X \cdot \frac{f(X)}{X} dX = \int_{X=0}^{\infty} f(X) dX. \quad (\text{Eq 105})$$

The relationship between the mean \bar{x} and the geometric mean μ_g and σ_g is:

$$\ln \bar{x} = \ln \mu_g + \sigma_g^2/2 \quad (\text{Eq 106})$$

For an adequate distribution of sample values A_{ai} from the same underlying measurement process, the parameters μ_g and σ_g^2 and thus $f(X)$ and the desired probability distributions can be obtained as follows:

1. Plot the cumulative fraction (or %) of A_{ai} values less than or equal to X versus selected values of X on a probability versus logarithmic scale paper, preferably marking the probabilities increasing in the upward direction and X values increasing to the right. Mark the vertical scale "% $\leq X$ " and the horizontal (logarithmically marked) as " X ".
2. Assuming that the distribution is log-normal, draw a best fit line through the points, emphasizing those in the 25% - 75% region, or fit the line by weighted regression techniques if the variances in A_{ai} vs. X can be estimated (61), (i.e.) or by the method of maximum likelihood, etc. (61).
3. Estimate the median μ_g as the X value where the line crosses the 50% horizontal median, or calculate it from:

$$\hat{\mu}_g = \left(\prod_{i=1}^N A_{ai} \right)^{1/N}, \quad (\text{Eq 107})$$

assuming N values of A_{ai} were obtained. (Taking the logarithm of Equation (107) to get $\ln \hat{\mu}_g$ would provide a maximum likelihood estimate of the mean of the $\ln A_{ai}$ distribution (51).

4. Estimate the "standard geometric deviation" (sometimes called "geometric mean standard deviation") from:

$$S_g = x_{0.8413}/x_{0.50} = x_{0.50}/x_{0.1587}, \quad (\text{Eq 108})$$

where $x_{0.8413}$, $x_{0.50}$, and $x_{0.1587}$ are the values of x where the log-normal line crosses the horizontal 84.13%, 50%, or 15.87% lines, respectively. Then, estimate the value of σ_g from:

$$\hat{\sigma}_g = \ln S_g. \quad (\text{Eq. 109})$$

NOTE: The derived relative bias $B_{ri} = \frac{A_{ai} - A_a}{A_a}$ (or $B_r = \bar{B}_{ri}$) can not be

itself log-normally distributed, or usefully plotted on log-normal probability paper, since it has negative values when measurements $A_{ai} \leq A_a$. Thus, when the data A_{ai} are log-normally distributed in a given way, the probabilities of finding B_r values within any range must be calculated as shown below.

3.9.4 Calculation of the Probability of \bar{B}_r Falling Within a Specified Range

From Section 2

$$\bar{B}_r = \frac{\sum_{i=1}^N B_{ri}}{N} = \frac{1}{N} \sum_{i=1}^N \frac{A_{ai} - A_a}{A_a} \quad (\text{Eq 110})$$

Our need is to calculate

$$P(b_1 \leq \bar{B}_r \leq b_2) = P(b_1 \leq \frac{1}{N} \sum_{i=1}^N \left(\frac{A_{ai} - A_a}{A_a} \right) \leq b_2) \quad (\text{Eq 111})$$

Now expression (111) may be written

$$P(N b_1 \leq \sum_{i=1}^N \left(\frac{A_{ai} - A_a}{A_a} \right) \leq N b_2) \quad (\text{Eq 112})$$

Case Where All A_a (actual spiked activities) Are the Same, or the B_{ri} Distributions Are The Same at Any Level of A_a

In this case, where there is a consistent underlying distribution in biases B_{ri} , the term A_a may be factored out of expression (Eq 112) as follows:

$$P(N b_1 \leq \frac{1}{A_a} \sum_{i=1}^N (A_{ai} - A_a) \leq N b_2), \quad (\text{Eq 113})$$

which is the same as

$$P(A_a N b_1 \leq \sum_{i=1}^N A_{ai} - N A_a \leq A_a N b_2), \quad (\text{Eq 114})$$

which is also the same as

$$P(N A_a (1 + b_1) \leq \sum_{i=1}^N A_{ai} \leq N A_a (1 + b_2)) \quad (\text{Eq 115})$$

Thus, if the probability of Equation 115 could be found, it would be the same as the probability in Equation 111 that we are seeking.

5. Calculations for Normally Distributed A_{ai} . If the A_{ai} are normally distributed with mean μ_A and variance σ_A^2 , then $\sum_{i=1}^N A_{ai}$ would be normally distributed with mean $N\mu_A$ and variance $\sigma_{NA}^2 = N\sigma^2$. Then, the probability of Equation (114) may be calculated by a transformation of the distribution of A_{ai} to a unit normal distribution and use of the standardized unit normal ($N(0,1)$) distribution tables. That is, find for the distribution parameters μ and σ^2 fit to a sample, the following probability in the standard unit normal tables:

$$P\left(\frac{NA_a(1 + b_1) - N\mu_A}{\sqrt{N\sigma^2}} \leq N(0, 1) \leq \frac{NA_a(1 + b_2) - N\mu_A}{\sqrt{N\sigma^2}}\right) \quad (\text{Eq 116})$$

This could of course be calculated and plotted versus the true bias of the measurements, $A_a - \mu_A$, and for various values of b_1 and b_2 as specified limits, for given sample size N .

6. Calculations for the Case of Log-Normal A_{ai} . In this case, $\sum_{i=1}^N A_{ai}$ would not ordinarily be log-normally distributed except for the special case $N = 1$.

If $N > 1$ and the distribution A_{ai} have been found to be log-normal, then the best approach to the calculation of (111) would be to determine the μ_g and σ_g parameters of $f(X)$ as described before, transform all A_{ai} to $\ln A_{ai} = \ln_i$ values and transform the limits b_1 and b_2 to corresponding limits a_1 and a_2 using the definition of relative bias,

$$B_{ri} = \frac{A_{ai} - A_a}{A_a}, \quad (\text{Eq 117})$$

i.e.,

$$a_1 = A_a b_1 + A_a = A_a (1 + b_1) \quad (\text{Eq 118})$$

and

$$a_2 = A_a b_2 + A_a = A_a (1 + b_2) \quad (\text{Eq 119})$$

Then, the desired probability of Eq (111) may be calculated using a relationship analogous to Equation (117). However, in the log-normal case, the geometric mean of the actual sample values should be a better (more unbiased) estimator of the underlying central value (median) of the measurements. Thus, the problem becomes analogous to expressions (111) to (117).

$$\hat{\mu}_A = \left(\prod_{i=1}^N A_{ai} \right)^{1/N} \quad (\text{Eq 120})$$

$$\ln \hat{\mu}_A = \frac{1}{N} \sum \ln A_{ai} \quad (\text{Eq 121})$$

Now $P(\ln a_1 \leq \ln \hat{\mu} \leq \ln a_2)$, (Eq 122)

where a_1 and a_2 are obtained from Equations (118) and (119), becomes successively:

$$P(\ln a_1 \leq \frac{1}{N} \sum \ln A_{ai} \leq \ln a_2) \quad (\text{Eq 123})$$

$$P(N \ln a_1 \leq \sum \ln A_{ai} \leq N \ln a_2) \quad (\text{Eq 124})$$

Now, $\sum_{i=1}^N \ln A_{ai}$ would have a normal distribution (since the A_{ai} are log-normally distributed) with mean

$$\ln M_g = N \ln \mu_g, \quad (\text{Eq 125})$$

and variance,

$$\sigma_a^2 = N\sigma_g^2 \quad (\text{Eq 126})$$

so that the desired probability of staying within the specified biases may be estimated from standardized unit normal distribution tables to obtain the probability that the unit normal $N(0,1)$ random variable will lie between limits as follows:

$$P\left(\frac{N \ln a_1 - N \ln \mu_g}{\sqrt{N\sigma_g^2}} \leq N(0,1) \leq \frac{N \ln a_2 - N \ln \mu_g}{\sqrt{N\sigma_g^2}}\right) \quad (\text{Eq 127})$$

NOTE: Of course, with any finite sample, the above analyses can give only approximate probabilities based on estimated parameters of the underlying probability distributions of error. If a measurement system remains the same for a series of test samples, then the parameters of the underlying distributions can be estimated more and more precisely, for obtaining better predictions of the probabilities of future sets of samples lying within any specified limits of bias.

3.10 Computer Program and Example Calculations

The computer program of Exhibit 2 was written for the HP 41CV computer to organize computations of $P(b_1 \leq B_r \leq b_2)$, and to avoid tedious calculations, interpolations, and subtractions of probabilities given in tables of cumulative probabilities under the standardized normal distribution. The program asks for the

necessary data, and may be calculated for any sample size N, any spiked value AA (A_a in text), and either normally or log-normally distributed A_{ai} . A trapezoidal integration under the standardized unit normal curve is included in the program, from a standardized lower limit LL to a standardized upper limit UL, corresponding to b_1 and b_2 , respectively. Although the interval size dX for numerical integration may be selected, the program provides the required probabilities to within less than 0.1% of those from normal tables using an interval of 0.01.

The two examples below illustrate the calculations of $P(b_1 \leq B_r \leq b_2)$.

Example for Normally Distributed A_{ai}

Assuming that a set of cumulative data from a laboratory repeatedly tested in a given analytical test category have been plotted on normal probability (versus linear) paper and found to be approximated by a straight line as in Figure 4, the parameters of a normal distribution of A_{ai} may be estimated from the graph as:

$$\begin{aligned} \mu &= \text{mean and median} = 5.25 \\ \sigma &= \text{standard deviation (SIG in program of Exhibit 2)} \\ &= 1.73 \end{aligned}$$

In the computer program of Exhibit 2, these values of m and s are entered as data. Also, the "true" value of the spike is assumed as $AA = 4.33$, $L = 0$ to represent a normal rather than log-normal distribution calculation, and an assumed sample size $N = 3$ for the next test run, as well as the relative biases specified as limits in the ANSI standard, $b_1 = -0.25$ and $b_2 = 0.50$. Also, for b_1 and b_2 the computer calculates $LL, N = -2.0049$ and $UL, N = 1.2465$ as the equivalent lower and upper limits of integration of the standardized unit normal, and these are entered as LL and UL, respectively.

The computer output is

$$P(b_1 \leq B_r \leq b_2) = 0.8728, \quad (\text{Eq 128})$$

equivalent to the expression of Equation (116), which becomes the same as

$$P(-2.0049 \leq z \leq 1.2465) = 0.8728, \quad (\text{Eq 129})$$

where z is the standardized unit normal variate.

This result may be interpreted as saying that with the underlying normal distribution of A_{ai} values of Figure 4, and a true spike AA, the probability is 87.3% that the bias B_r calculated from a sample of 3 spikes each having actually, e.g., $4.33 \mu\text{g/liter}$, will fall within $(-0.25, + 0.50)$.

Example of log-normally distributed A_{ai}

For illustration, consider the example data and log-normal distribution of A_{ai} in Figure 2, plotted as described in Section 3.9.3. The parameters taken from this graph are:

Median = 5.25 (not now equal to the mean)
SIG = 1.73

These values and the following are taken to be the same as in the previous example, for comparison purposes only:

AA = 4.33
 $b_1 = -0.25$; $b_2 = + 0.50$
N = 3

Now $L = 1$ is selected in the computer program, and the computer program gives only the LL, LN = -0.4809 and UL, LN = 0.2131 values for the log-normal distribution integration, which are put in as data for LL and UL, respectively. The result is:

$$P(b_1 \leq B_r \leq b_2) = P(LL = - 0.4809 < Z < 0.2131 = UL) \quad (\text{Eq 130}) \\ = 0.2714, \text{ or } 27.1\%.$$

The main reason that a much smaller chance is obtained for the B_r falling within the bias criteria interval (b_1, b_2) for this log-normal distribution is that the value of $\sigma = \text{SIG}$ for this distribution, = 1.73, is the standard deviation of $\ln A_{ai}$, not A_{ai} . The "standard geometric deviation," S_g is calculated from σ and its meaning is illustrated in Figure 5. Considering the logarithmic scale, and the interval of (0.931 to 29.6) now spanning most of the data, it is clear that the log-normal distribution of Example 2 is a much wider distribution than that of Example 1, so that more of the A_{ai} values will lie outside the (b_1, b_2) bias interval.

3.11 Example of Calculation of Probability of Passing Both the Bias and Precision Tests for N = 3

Consider just the estimators of bias, B_r , and precision, S_B , of measurement when they are independent stochastically, i.e., when the bias is an "additive" error - a constant error in a term added (or subtracted) to obtain the measurement result. The probability of passing both the bias and precision tests using parameters from Example 1, and interpolating for $\sigma/a = 1.73/4.33 = 0.3995$ and $N = 3$ from Table 2, is:

$$P(\text{passing bias and precision tests}) = 0.8728 \times (0.63) = 0.55, \quad (\text{Eq 131})$$

or 55%, even though the "true" precision is about 0.33 (<0.4) and the true bias $\frac{5.25-4.33}{4.33} = 0.2125$, both within the ranges for S_B and B_r .

The probability of passing six test categories with one triplicate sample in each category would thus be $(0.55)^6 = 0.028$, if relative bias and precision conditions were the same for each category. If two sample test levels were

used in each of six categories, then the probability of passing all tests would be: $0.028 \times 0.028 = 0.000784$, only about 8 chances in 10,000.

Table 4 presents the true underlying relative biases and precisions of measurement required to obtain the probabilities shown in the table of passing 6 and 12 test categories, each with $N=3$ replicates, for the Pass/Fail criteria $|S_B| \leq 0.4$, $-0.25 \leq B_r \leq 0.50$, for each test category triplicate measurement.

Table 4
Probabilities of Passing Bias and Precision Tests

(Requirements of true precision and bias to pass multi-category tests with only two pass/fail criteria of $|S_B|$ or $|S_A| \leq 0.4$ and $-0.25 \leq B_r \leq 0.50$, when these statistics are stochastically independent.)

Probabilities	True Relative Bias and Precision			
	B=+10%, P=10%	B=+15%, P=15%	B=-20%, P=20%	B=+20%, P=20%
P_S - passing precision test, one N = 3 category	1.0000	0.995466	0.997739	0.937741
P_B - passing bias test, one N = 3 category	0.9999	0.999756	0.7516	0.984503
$P_S P_B$ - passing bias and precision, one category	1	0.995224	0.7499	0.923210
$(P_S P_B)^6$ - passing 6 categories	1	0.97168	0.17784	0.619159
$(P_S P_B)^{12}$ - passing 6 categories at 2 levels each, or 12 categories	1	0.944	0.031626	0.383358

The same Table 4 values would be obtained when constant (bias) errors occur in multiplicative terms only, and the two criteria are $|S_A| \leq 0.4$ and $-0.25 \leq B_r \leq 0.50$; however, the % values of P in the column headings are then interpreted as the underlying coefficients of variation, $E(S_A/A)$. When bias errors are mixed, or when the pair of estimators for bias and precision are not independent, then the values in Table 4 become upper limits of the probabilities of passing or failing the pair of criteria. Also, when both precision criteria, $|S_B|$ and $|S_A|$ both ≤ 0.4 , as well as the bias criteria are invoked, as now proposed for the draft ANSI N13.30 standard (24), then the probabilities of passing are again somewhat less than the values in Table 4. Thus, Table 4 can give only the best possible chances of passing the bias and precision criteria in the general case.

The exact probabilities of passing selected performance criteria for sample measurements with small numbers of replicates can be determined for any real interlaboratory testing program only after a large amount of data are obtained for an ongoing testing program under consistent conditions over time. Thus, the analytical methods presented in this section are useful only to give approximate estimates of the underlying degrees of accuracy necessary in analytical procedures to pass various accuracy criteria as calculated based on objective sample statistics.

EXHIBIT 2 - Program for Computing $P(b_1 \leq B_r \leq b_2)$ for Either Normally or Lognormally Distributed Sample Measurements

```

001 *LOGNRM*
02 *THIS PROGRAM*
03 *AVIEW*
04 *CALCULATES THE*
05 *PROBABILITY OF*
06 *PASSING THE*
07 *TENS TEST*
08 *FIX 4*
09 *MEDIAN=?*
10 *SIG=?*
11 *B1=?*
12 *B2=?*
13 *L=0 OR 1?*
14 *N=?*
15 *R=?*
16 *CALC. INT. LIMITS*
17 *AVIEW*
18 *PSE*
19 *RCL 02*
20 *ENTER*
21 *RCL 06*
22 *RCL 06*
23 *RCL 06*
24 *RCL 06*
25 *RCL 06*
26 *RCL 06*
27 *RCL 06*
28 *RCL 06*
29 *RCL 06*
30 *RCL 06*
31 *RCL 06*
32 *RCL 06*
33 *RCL 06*
34 *RCL 06*
35 *RCL 06*
36 *RCL 06*
37 *RCL 06*
38 *RCL 06*
39 *RCL 06*
40 *RCL 06*
41 *RCL 06*
42 *RCL 06*
43 *RCL 06*
44 *RCL 06*
45 *RCL 06*
46 *RCL 06*
47 *RCL 06*
48 *RCL 06*
49 *RCL 06*
50 *RCL 06*
51 *RCL 06*
52 *ENTER*
53 *
54 *
55 *RCL 07*
56 *
57 *STO 10*
58 *RCL 05*
59 *X=0?
60 *GTO 01*
61 *GTO 02*
62 *LBL 01*
63 *RCL 09*
64 *ENTER*
65 *RCL 01*
66 *
67 *RCL 06*
68 *
69 *RCL 08*
70 *
71 *STO 11*
72 *RCL 10*
73 *ENTER*
74 *RCL 01*
75 *
76 *RCL 06*
77 *
78 *RCL 08*
79 *
80 *STO 12*
81 *LBL 02*
82 *RCL 09*
83 *LN*
84 *ENTER*
85 *RCL 01*
86 *LN*
87 *
88 *RCL 06*
89 *
90 *RCL 08*
91 *
92 *STO 13*
93 *RCL 10*
94 *LN*
95 *ENTER*
96 *RCL 01*
97 *LN*
98 *
99 *RCL 06*
100 *
101 *RCL 02*
102 *
103 *STO 14*
104 *RCL 11*
105 *LOWER LIM 2*
106 *AVIEW*
107 *OF INTEGRATION*
108 *AVIEW*
109 *FOR Z NORMAL*
110 *AVIEW*
111 *DISTN OF A*
112 *AVIEW*
113 *PSE*
114 *LL,N=*
115 *RCL X*
116 *AVIEW*
117 *PSE*
118 *PSE*
119 *PSE*
120 *PSE*
121 *PSE*
122 *PSE*
123 *RCL 12*
124 *UPPER LIM 2*
125 *AVIEW*
126 *NORMAL A*
127 *AVIEW*
128 *PSE*
129 *UL,N=*
130 *RCL X*
131 *AVIEW*
132 *PSE*
133 *PSE*
134 *PSE*
135 *PSE*
136 *PSE*
137 *PSE*
138 *X=0?
139 *GTO 04*
140 *GTO 03*
141 *LBL 03*
142 *RCL 13*
143 *LOWER LIM 2*
144 *AVIEW*
145 *LOGNORMAL A*
146 *AVIEW*
147 *PSE*
148 *LL,N=*
149 *RCL X*
150 *AVIEW*
151 *PSE*
152 *PSE*
153 *PSE*
154 *PSE*
155 *PSE*
156 *PSE*
157 *RCL 14*
158 *UPPER LIM 2*
159 *AVIEW*
160 *LOGNORMAL A*
161 *AVIEW*
162 *PSE*
163 *UL,N=*
164 *RCL X*
165 *AVIEW*
166 *PSE*
167 *PSE*
168 *PSE*
169 *PSE*
170 *PSE*
171 *PSE*
172 *SELECT LL*
173 *AVIEW*
174 *PSE*
175 *LL=?*
176 *PROMPT*
177 *STO 15*
178 *SELECT UL*
179 *AVIEW*
180 *PSE*
181 *UL=?*
182 *PROMPT*
183 *STO 16*
184 *RCL 15*
185 *STO 17*
186 *RCL 16*
187 *STO 18*
188 *INCREMENT*
189 *AVIEW*
190 *INC=.01 OR .001*
191 *PROMPT*
192 *STO 19*
193 *
194 *STO 22*
195 *LBL 04*
196 *RCL 17*
197 *ENTER*
198 *
199 *
200 *
201 *CHK*
202 *E+X*
203 *STO 20*
204 *RCL 17*
205 *ENTER*
206 *RCL 19*
207 *
208 *ENTER*
209 *
210 *
211 *
212 *CHK*
213 *E+X*
214 *ENTER*
215 *RCL 20*
216 *
217 *
218 *
219 *STO 21*
220 *
221 *ENTER*
222 *3.1416*
223 *
224 *SOFT*
225 *1/X*
226 *ENTER*
227 *RCL 21*
228 *
229 *ENTER*
230 *RCL 19*
231 *
232 *ST+ 22*
233 *RCL 19*
234 *ENTER*
235 *RCL 17*
236 *
237 *STO 17*
238 *RCL 18*
239 *X/Y?
240 *GTO 04*
241 *GTO 05*
242 *LBL 05*
243 *RCL 22*
244 *PLOCKUL*
245 *RCL X*
246 *AVIEW*
247 *PSE*
248 *PSE*
249 *PSE*
250 *END.

```

FINDING PARAMETERS FOR NORMAL A_{ai}

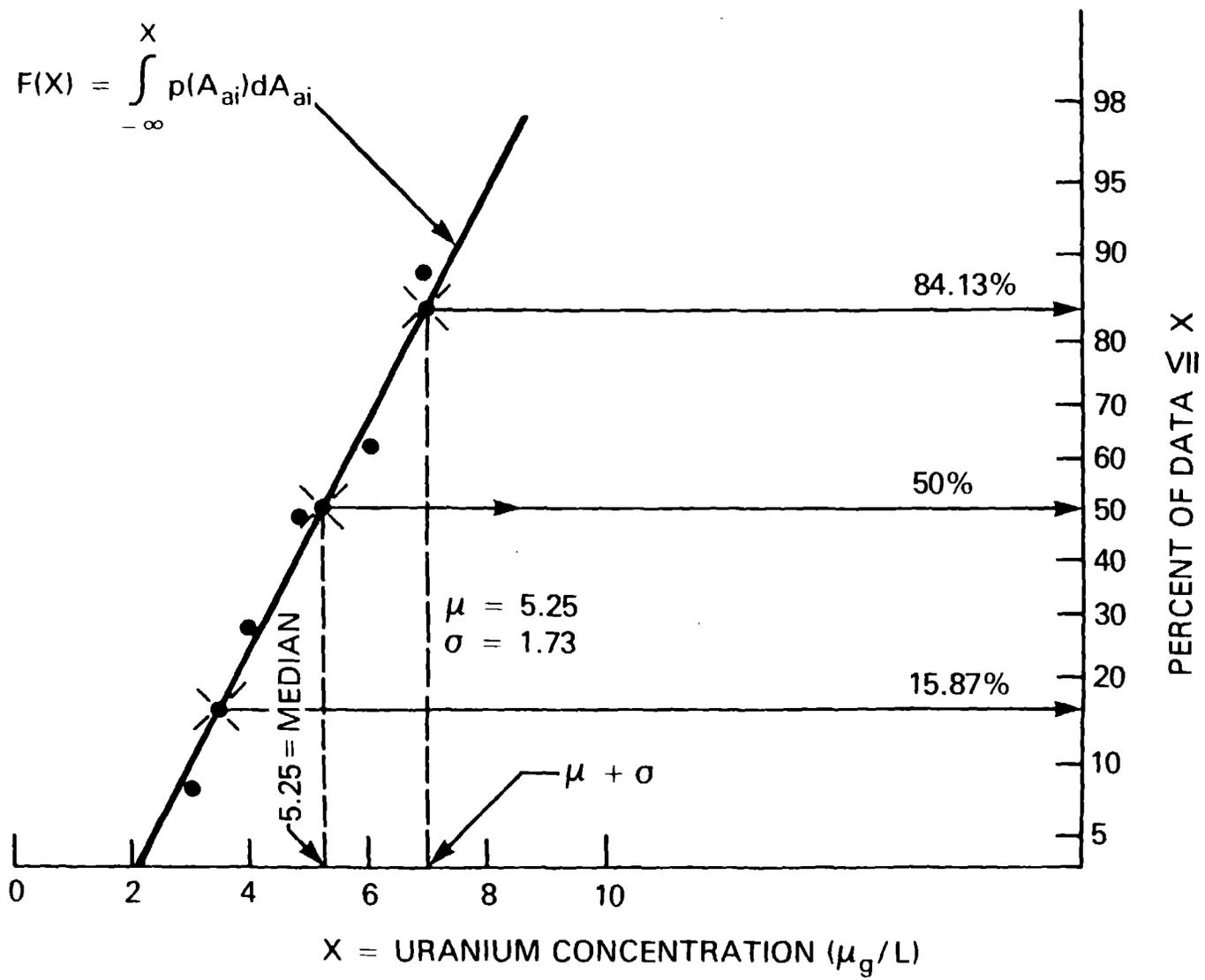


FIGURE 4

FINDING PARAMETERS FOR LOGNORMAL A_{ai}

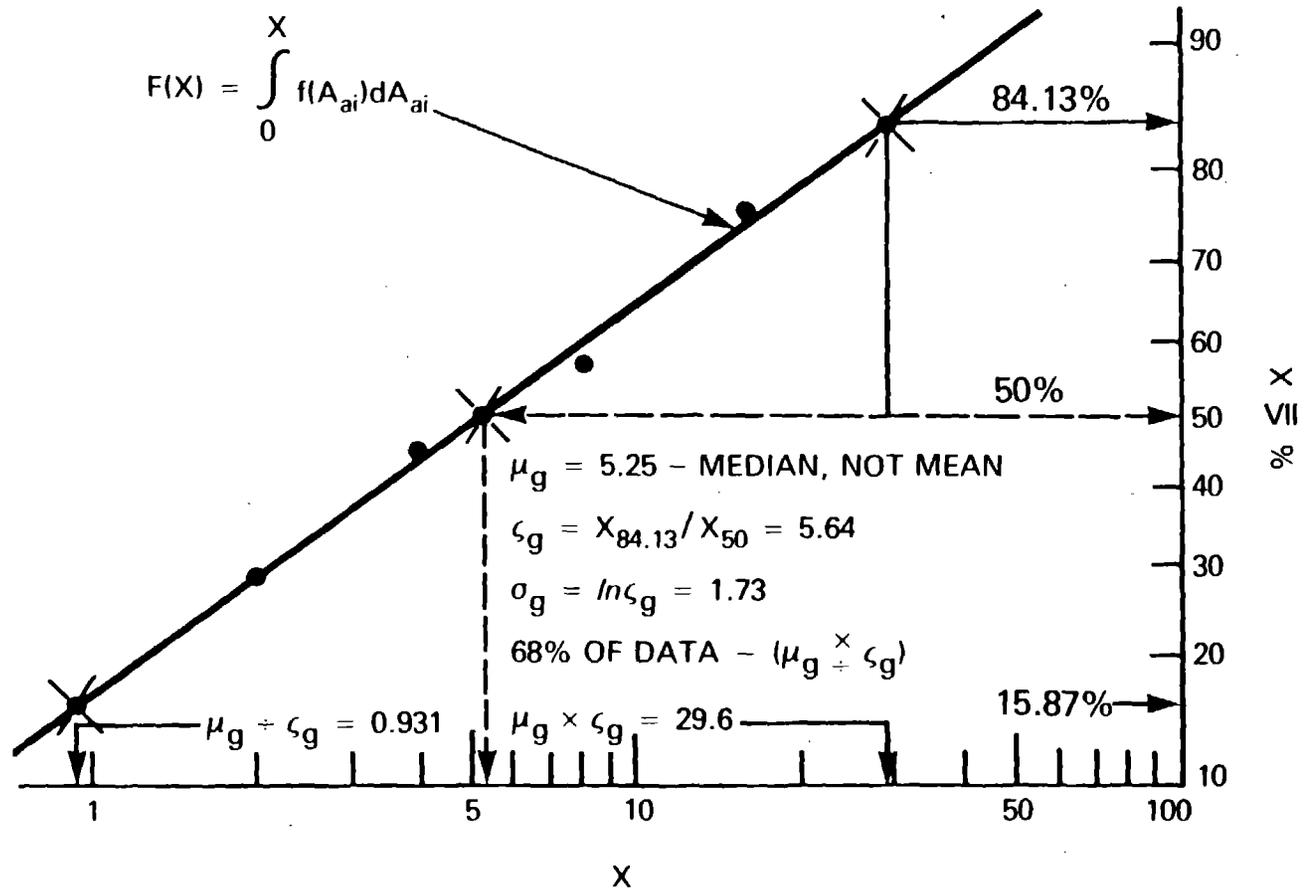


FIGURE 5

4 DISCUSSION

The formulations presented in this report for expressing MDA, bias and precision are not the only possible formulations, and are not necessarily consistent with those used by other capable analytical laboratories or testing programs. They have been suggested for use in developing performance criteria for radio-bioassay laboratories, for the reasons discussed along with their respective presentations. The MDA formulations in this report would provide greater assurance that the stated MDA quantity would actually be detected than the use of criteria such as two or three times the standard deviation in the "background count." (127) The bias and precision statistics, B_r and S_B presented in this report turn out to be the same sample statistics as B_r adopted in the ANSI N13.11 standard on personnel dosimetry performance (128), except that these two statistics are added together to obtain a single performance statistic for the ANSI N13.11 standard, but are used separately in this report and in the draft ANSI N13.30 standard to separately evaluate laboratory performance errors due to fixed biases and random fluctuations of measurement. The use of these two estimators separately has now also been recommended by persons who participated in developing the ANSI N13.11 standard. (129, 130) The question of combining random and fixed errors into an overall accuracy statement is a topic of continuing scientific discussion (131), and one that need not prevent the use of separate (and sometimes independent) statistical estimators for assessing the performance of separate aspects of laboratory analyses of test samples.

The Working Group (24) and this author have also preferred to base accuracy criteria for small samples on simple formulas for the bias and precision estimators, rather than on parameters related to standard distribution statistics such as the T or chi-square (132), which depend on the assumption of Gaussian errors. Analyses have been presented in this document, with algorithms, for calculating probabilities of passing any number of replicate tests, using the statistics B_r and S_B , or B_r and S_A , as appropriate. The use of both S_B and S_A in testing precision performance complicates the probability analysis somewhat, but always in a way in which the maximum probability of passing test criteria can be evaluated. Whatever the test statistic used for expressing precision of measurements, it should be applicable to random errors having coefficients of variation of 50% or more, and possibly skewed error distributions. (133)

As indicated in the preface, although it is expected that this report will be useful in radioanalytical quality assurance standardization or programming, the author does not intend or expect that it provides the last word on the subject. Any comments or suggestions, and further literature references, would be appreciated by the author if called to his attention.

REFERENCES

1. BUCHANAN, J. D. Sensitivity and Detection Limit, Health Physics, Vol. 33, 1977, pp. 347-348.
2. BRODSKY, A. and KATHREN, R. L. Accuracy and Sensitivity of Film Measurements of Gamma Radiation - Part I, Comparison of Multiple Film and Single-Quarterly-Film Measurements of Gamma Dose at Several Environmental Conditions, Health Physics, Vol. 9, 1963, pp. 453-462.
3. BRODSKY, A. Accuracy and Sensitivity of Film Measurements of Gamma Radiation - Part II, Limits of Sensitivity and Precision, Health Physics, Vol. 9, 1963, pp. 463-472.
4. KATZ, L. Presentation of a Confidence Interval Estimate As Evidence in a Legal Proceeding, The American Statistician, Vol. 29, 1975, pp. 138-142.
5. HEALY, J. W., talk presented in Proceedings of the Second Annual Meeting on Bioassay, U.S.A.E.C. Report WASH-736, September 1957.
6. ALTSHULER, B. and PASTERNAK, B., Statistical Measures of the Lower Limit of Detection of a Radioactivity Counter, Health Physics, Vol. 9, 1963, pp. 293-298.
7. CURRIE, L.A. Limits for Qualitative Detection and Quantitative Determination, Analytical Chemistry, Vol. 40, No. 3, 1968, pp. 586-593.
8. PASTERNAK, B. S. and HARLEY, N. H. Detection Limits for Radionuclides in the Analysis of Multi-Component Gamma Ray Spectrometer Data, Nuclear Instruments and Methods, Vol. 91, 1971, pp. 533-40.
9. FISENNE, I., O'TOOLE, A., and CUTLER, R. Least Squares Analysis and Minimum Detection Levels Applied to Multi-Component Alpha Emitting Samples, Radiochem. Radioanal. Letters, 16/1, 1973, pp. 5-16.
10. HARTWELL, J. K. Detection Limits for Radioanalytical Techniques, ARH-SA-215, Atlantic Richfield Company, Richland, WA 99352, 1975.
11. HARLEY, J. H., Editor "HASL Procedures Manual, HASL-300 (Suppl. 4), Technical Information Center, Energy Research and Development Admin., Washington, DC, 1976, pp. D-08-01 to D-08-12.
12. LOCHAMY, J. C. The Minimum-Detectable-Activity Concept, in NBS Special Publication 456, Measurements for the Safe Use of Radiation, U.S. Department of Commerce, National Bureau of Standards, Washington, DC, 1976.
13. DONN, J. J. and WOLKE, R. L. The Statistical Interpretation of Counting Data from Measurements of Low-Level Radioactivity, Health Physics, Vol. 32, 1977, pp. 1-14.

14. Draft NRC Staff Specifications, Radiological Effluent Technical Specifications for Pressurized Water Reactors, NUREG-0472, Rev. 3, Office of Nuclear Reactor Regulation, U.S. Nuclear Regulatory Commission, Washington, DC 20555, March 1979 (see January 1983 revision).
15. Regulatory Guide 4.14, Measuring, Evaluating, and Reporting Radioactivity in Releases of Radioactive Materials in Liquid and Airborne Effluents from Uranium Mills, U.S. NRC, Washington, DC, 1980.
16. Regulatory Guide 4.15, Quality Assurance for Radiological Monitoring Programs (Normal Operations) - Effluent Streams and the Environment, U.S. NRC, Washington, DC, 1979.
17. NICHOLSON, W. L., Nucleonics, Vol 24, 1966, p. 118.
18. BARRY, B. J. Comments on "The Statistical Interpretation of Counting Data from Measurements of Low-Level Radioactivity" by J. J. Donn and R. L. Wolke, Letter, Health Physics Vol 36, 1979, pp. 208-209.
19. American National Standards Institute, ANSI N15.5-1982, Statistical Terminology and Notation for Nuclear Materials Management, American National Standards Institute, Inc., 1430 Broadway, New York; NY 10018.
20. CURRIE, L. A., Lower Limit of Detection: Definition and Elaboration of a Proposed Position for Radiological Effluent and Environmental Measurements, NUREG/CR-4007, National Bureau of Standards, September 1984.
21. GINEVAN, M. E. A Poisson Trials Approach to Interpopulation Comparisons of Cause of Death Data, Environmental Research, Vol 25, 1981, pp. 147-159.
22. ROBINSON, A. V., FISHER, D. R., and HADLEY, R. T. Performance Testing of Radiobioassay Laboratories: In Vitro Measurements, Pilot Study Report, NUREG/CR-3809, Vol 1, (DOE/NBM-1071), DOE-NRC sponsored report, December 1984.
23. REESE, W. D. et al., Performance Testing of Radiobioassay Laboratories: In Vivo Measurements Pilot Study Report. DOE-4840, U.S. Department of Energy, Washington, DC 20545, in press, 1986.
24. HEID, K. R. Chairman. ANSI Working Group N13.30. Draft Standard, Performance Criteria for Radiobioassay, 1985 revision. Battelle-Pacific Northwest Laboratories, P.O. Box 999, Richland, WA 99352.
25. American National Standards Institute, ANSI N42.14-1978, Calibration and Usage of Germanium Detectors for Measurement of Gamma-Ray Emission of Radionuclides, American National Standards Institute, 1430 Broadway, New York, NY 10018 (published by Institute of Electrical and Electronics Engineers, Inc., 345 East 47th Street, New York, NY 10017).
26. WAGNER, S. How to Treat Systematic Errors in Order to State the Uncertainty of a Measurement. Bericht FMRB 31/69, Physikalisch-Technische Bundesanstalt Forschungs-und MeBreaktor Braunschweig, November 1969.

27. EISENHART, C. Expression of the Uncertainties of Final Results. *Science*, vol 160, p 1202, June 1968.
28. International Commission on Radiation Units and Measurements. Certification of Standardized Radioactive Sources, ICRU Report No 12, September 15, 1968.
29. GOTCHY, R. L. and BORES, R. J. The Public Whole Body Counting Program Following the Three Mile Island Accident. NUREG-0636, U.S. Nuclear Regulatory Commission, Washington, DC 20555, December 1980.
30. LIUZZI, A. and PASTERNAK, B. S. Nuclear Instruments and Methods vol 57, p 229, 1967.
31. ROGERS, W. J. Lower Limit of Detection Calculations for Sr-89 and Sr-90 Determined Simultaneously. Laboratory Services Report No. 19, Tennessee Valley Authority, Radiological Health Staff, Western Area Radiological Laboratory - M, Muscle Shoals, Alabama 35660, June 1984.
32. BOWMAN, W. B., SWINDLE, H. D. L. and HENDERSON, D. L. Radiochemical Determination of ⁸⁹Sr AND ⁹⁰Sr: A Comparison of Methods Based on Error Analysis. Health Physics vol 31, pp 495-500, 1976.
33. BRODSKY, A. Statistical Methods of Data Analysis. Chapter 6.2 in Handbook of Radiation Measurement and Protection, Section A, Volume II, Biological and Mathematical Information, edited by A. Brodsky, Boca Raton, FL: CRC Press, Inc., 1982, pp 317-322.
34. SEITZ, F.; and MUELLER, D. On the Statistics of Luminescent Counter Systems. Physical Review, vol 78, 1950, pp 605-613.
35. CRAWFORD-BROWN, D. J.; and WILSON, J. Observation on Very Long-Term Removal of Uranium Compunds. Health Physics, vol 47, no. 3, 1984, pp 443-446.
36. PALMER, H. E.; and RIEKSTS, G. The Use of Planar High-Purity Ge Detectors for In Vivo Measurement of Low-Energy Photon Emitters. Health Physics, vol 47, no. 4, October 1984, pp 569-578.
37. FALK, R. B.; TYREE, W. H.; WOOD, C. B.; and LAGERQUIST, C. R. A System of High Purity Germanium Detectors for the Detection and Measurement of Inhaled Radionuclides, in: IAEA Symposium on Advances in Radiation Protection Monitoring, Vienna: IAEA, 1978, pp 445-455.
38. YANIV, S. Plutonium and Americium Measurement in Humans by X- and Gamma-Ray Spectral Analysis. Report WASH-1241, Washington, DC: U. S. Atomic Energy Commission, 1973 (available from U. S. Government Printing Office, Washington, DC 20402).
39. VASILIK, D. G.; AIKIN, I.; COOP, K. L.; and UMBARGER, C. J. In Vivo Assessment of Lung Burdens at the Los Alamos National Laboratory. Report LA-9979-MS, Los Alamos: Los Alamos National Laboratory, Los Alamos, New Mexico 87545, 1984.

40. INTERNATIONAL COMMISSION ON RADIOLOGICAL PROTECTION. Recommendations of the International Commission on Radiological Protection, report of Committee II on permissible dose for internal radiation. W. S. Snyder, Chairman. ICRP Publication 2. Health Physics, vol 3, 1960.
41. INTERNATIONAL COMMISSION ON RADIOLOGICAL PROTECTION. Limits for Intakes of Radionuclides by Workers, ICRP Publication 30 and supplements. Oxford: Pergamon Press, 1978.
42. MORAN, M. G.; and KOWALSKI, B. R. Effect of Random Experimental Error on the Generalized Standard Addition Method. Analytical Chemistry, vol 56, 1984, pp 562-569.
43. STROM, D. J.; GRAWFORD-BROWN, D. J.; and TANKERSLEY, W. G. Preparation of 30 Years' Personnel Monitoring Data from a Uranium Fabrication Plant for Use in Epidemiologic Studies. Presented at the 29th Annual Meeting of the Health Physics Society, available from Dr. D. J. Strom, Department of Radiation Health, University of Pittsburgh, Pittsburgh, PA 15261.
44. BRODSKY, A.; SPRITZER, A. A.; FEAGIN, F. E.; BRADLEY, F. J.; KARCHES, G. J.; and MANDELBERG, H. I. Accuracy and Sensitivity of Film Measurements of Gamma Radiation - Part IV, Intrinsic and Extrinsic Errors. Health Physics, vol 11, 1965, pp 1071-1082.
45. BRODSKY, A.; McNIGHT, A.; WALD, N.; and CATLIN, R. J. Cumulative Occupational Exposure Distributions in an Atomic Energy Facility. Health Physics, vol 10, 1964, p 613.
46. BOOTH, L. F.; and BRONSON, F. Internal Dosimetry and Record Keeping Implications of the Revisions to 10 CFR 20. Prepared for the Atomic Industrial Forum, Inc., National Environmental Studies Project. Available from Radiation Management Corporation, 3356 Commercial Avenue, Northbrook, IL 60062, 1984.
47. CURRIE, L.; GERLACH, R. W.; KLOUDA, G. A.; RUEGG, F. C. and TOMPKINS, G. B. Miniature Signals and Miniature Counters: Accuracy Assurance via Microprocessors and Multiparameter Control Techniques. Radiocarbon, Vol 25, No. 2, 1983, pp 553-564.
48. GOOD, W. M. and GOANS, R. E. Detection of Internally Deposited Actinides, Part I: Studies on the ORNL Phoswich System in Operational Health Physics, Health Physics Society Ninth Midyear Topical Symposium, Compiled and edited by Paul L. Carson, William R. Hendee, and D. C. Hunt, Denver, Central Rocky Mountain Chapter, Health Physics Society, P.O. Box 3229, Boulder, Colorado 80303, Feb. 9-12, 1976, pp 562-567.
49. GOANS, R. E. Detection of Internally Deposited Actinides, Part II: Statistical Techniques and Risk Analysis, in Operational Health Physics, *ibid*, pp. 568-573.
50. McCURDY, D. E. and MELLOR, R. A. ANSI Draft Standard N13.30 Comments, Memorandum to ANSI N13.30 Committee members. Yankee Atomic Electric Company, 1671 Worcester Road, Framingham, MA 01701, Aug. 8, 1985.

51. AITCHISON, J. and BROWN, J.A.C. The Lognormal Distribution, London: Cambridge at the University Press, 1963, p 1-50 and 154-155.
52. SCHUBERT, J.; BRODSKY, A.; and TYLER, S. The Log-Normal Function as a Stochastic Model of the Distribution of Strontium-90 and Other Fission Products in Humans. Health Physics, Vol 13, 1967, pp 1187-1204.
53. BRODSKY, A. Occupational Exposure Measurements in NRC Regulatory Guides. Published in Proceedings of a Meeting on Traceability for Ionizing Radiation Measurements. HEATON, H. T. II, Editor, NBS Special Publication 609, Center for Radiation Research, National Measurements Laboratory, National Bureau of Standards, U. S. Department of Commerce. Washington, D.C.; Government Printing Office, 1982.
54. INN, K. W. G. and NOYCE, J. R. The National Bureau of Standards Low-Level Radioactivity-Measurements Program. in HEATON, H. T. II, Editor, op. cit.
55. INTERNATIONAL COMMISSION ON RADIOLOGICAL PROTECTION, General Principles of Monitoring for Radiation Protection of Workers. ICRP Publication 12. Oxford, England: Pergamon Press, 1969.
56. NATIONAL COUNCIL ON RADIATION PROTECTION AND MEASUREMENTS (NCRP). Instrumentation and Monitoring Methods for Radiation Protection. NCRP Report No. 57. NCRP Publications, P.O. Box 30175, Washington, DC 20014, 1978.
57. SINCLAIR, W. K. The Scientific Basis for Risk Quantification. Published in the Proceedings of the Annual Meeting of the National Council on Radiation Protection and Measurements. Washington, DC: NCRP, 1980.
58. INTERNATIONAL COMMISSION ON RADIOLOGICAL PROTECTION. ICRP PUBLICATION 26, Recommendations of the International Commission on Radiological Protection. Oxford, England: Pergamon Press, 1977.
59. BRODSKY, A. Characteristics of Radiation Monitoring Instrumentation and Methods. Published in Medical Radiation Information for Litigation, edited by: Bushong, S.C., Cox, J.S.; Collins, V.P.; Neibel, J.B.; and Murphy, G.B. U.S. Department of Health, Education and Welfare Document No. DMRE 69-3. Washington, DC: U.S. Government Printing Office, July 1969.
60. BRADLEY, F.J.; WATSON, J.A.; DOOLITTLE, D.P.; BRODSKY, A; and SUTTON, R.B. RBE of 440-MeV Proton Radiation Compared to 125 Kvp X-rays for LD 50(30) of Mice. Health Physics, Vol 10, 1964, pp. 71-74.
61. FINNEY, D J. Probit Analysis, 2nd Ed. London: Cambridge University Press, 1962.
62. BRODSKY, A. A Stochastic Model of Carcinogenesis Incorporating Certain Observations from Chemical and Radiation Dose-Response Data. Health Physics, Vol 35, 1978, pp 421-428.

63. EVANS, R. D. Radium in Man. Health Physics, Vol 27, 1974, pp 497-510.
64. BRODSKY, A. Determining Industrial Hygiene Requirements for Installations Using Radioactive Materials. American Industrial Hygiene Journal, Vol 26, May-June 1965, pp 294-310.
65. POCHIN, E. E. Bases for the Detection and Measurement of Deposited Radionuclides. Published in Diagnosis and Treatment of Deposited Radionuclides, Edited by Korenberg, H.A.; and Norwood, W.D. New York: Excerpta Medica Foundation, 1968, pp 575-586.
66. BRODSKY, A. Accuracy and Sensitivity of Film Measurements of Gamma Radiation, Part II, Limits of Sensitivity and Precision. Health Physics, Vol 9, 1963, pp 463-471.
67. GOLDSTEIN, A. Biostatistics, An Introductory Text, New York; The MacMillan Co., 1971, pp 139-143.
68. BRODSKY, A.; SAYEG, J.A.; WALD, N.; WESCHLER, R. and CALDWELL, R. The Measurement and Management of Insoluble Plutonium and Americium Inhalation in Man. Published in Radiation Protection, Proceedings of the First International Congress of Radiation Protection, Part 2, Rome, Italy, Edited by Snyder, W.S.; Abee, H.H., Burton, L.K.; Maushart, R.; Benco, A; Oxford, England: Pergamon Press, 1968, pp 1181-1190.
69. WALD, N.; WECHSLER, R.; BRODSKY, A. and YANIV, S. Problems in Independent Medical Management of Plutonium-Americium Contaminated Patients. Published in Diagnosis and Treatment of Deposited Radionuclides, Edited by Kornberg H.A.; and Norwood, W.D. New York: Excerpta Medica Foundation, 1968, pp 575-586.
70. SNYDER, W. S. A Method of Interpreting Excretion Data Which Allows for Statistical Fluctuations of the Data. Published in Assessment of Radioactive Contamination in Man, STI/PUB/290. Vienna: International Atomic Energy Agency, 1972, pp 485-494.
71. SWINT, M. J.; and HEID, K. R. United States Transuranium Registry Annual Report, October 1, 1982 through September 30, 1983. HEHF 43. Richland, Washington, U.S.A.: Hanford Environmental Health Foundation, Richland, Washington 99352, April 1984.
72. AMERICAN NATIONAL STANDARDS INSTITUTE. USAS N13.3-1969 (Revised 1982), Guide for Administrative Practices in Radiation Monitoring. American National Standards Institute, 1430 Broadway, New York, N.Y. 10018, 1982.
73. AMERICAN NATIONAL STANDARDS INSTITUTE. ANSI N13.6-1966 (R 1972), American National Standard Practice for Occupational Radiation Exposure Record Systems. American National Standards Institute, 1430 Broadway, New York, N.Y. 1972.
74. AMERICAN NATIONAL STANDARDS INSTITUTE. ANSI N13.14-1963, American National Standard for Dosimetry - Internal Dosimetry Programs for Tritium Exposure - Minimum Requirements. American National Standards Institute, 1430 Broadway, New York, N.Y. 1983.

75. ICRP Publication 12, op. cit., p. 17.
76. INTERNATIONAL COMMISSION ON RADIOLOGICAL PROTECTION. ICRP PUBLICATION 35, General Principles of Monitoring for Radiation Protection of Workers. Oxford, England: Pergamon Press, 1983.
77. AMERICAN NATIONAL STANDARDS INSTITUTE. ANSI N343-1978, American National Standard for Internal Dosimetry for Mixed Fission and Activation Products. American National Standards Institute, 1430 Broadway, New York, N.Y. 10018.
78. INTERNATIONAL ATOMIC ENERGY AGENCY. Safety Series No. 14, IAEA Safety Standards, Basic Requirements for Personnel Monitoring. Vienna, Austria: International Atomic Energy Agency, 1980.
79. INTERNATIONAL ATOMIC ENERGY AGENCY. Safety Series No. 49, Procedures and Data, Radiological Surveillance of Airborne Contaminants in the Working Environment. Vienna, Austria: International Atomic Energy Agency, 1979.
80. SNYDER, W. S.; COOK, M. J.; KARHAUSEN, L. R.; NASSET, E. S.; HOWELLS, G. P.; and TIPTON, I. H. Report of the Task Group on Reference Man, ICRP Report No. 23, International Commission on Radiological Protection. Oxford, England: Pergamon Press, 1975.
81. IAEA Safety Series 49, Reference 79, op cit., p. 99, p. 18, pp. 47-59.
82. TRICE, C. Radiochemical Determination of Plutonium in Glass Fiber Filters. RFP-3653, UC-37, DOE/TIC-4500 (Rev. 72), Rockwell International, Energy Systems Group, Rocky Flats Plant, P.O. Box 464, Golden, Colorado 80402-0464, 1984.
83. WESSMAN, R. A. An Overview of the Radiochemical Analysis of Uranium. Presented in Biokinetics and Analysis of Uranium in Man, sponsored jointly by The United States Uranium Registry and The Joint Center for Graduate Study, Hanford Environmental Health Foundation, Richland, WA 99352, 1984.
84. SEDLET, J., FAIRMAN, W. D., AND ROBINSON, J. J. Rapid Methods of Determining Internal Radioactive Contamination, in Personnel Dosimetry for Radiation Accidents. Vienna: IAEA, 1965.
85. SEDLET, J. Analytical Methods Under Emergency Conditions, in Internal Radiation Dosimetry, Edited by Brodsky, A., Mossman, K., Hering, H., McCormick, W. E., and Fisher, D. R., Stow, Ohio: Envirotex Management, Inc., in press, 1986.
86. Code of Federal Regulations, Title 10 Chapter 1, Part 20, Standards for Protection Against Radiation, U.S. Government Printing Office, 1984.
87. United States Department of Energy, Manual Chapter 0524, "Standards for Radiation Protection," and Manual Chapter 0525, "Occupational Radiation Exposure Information," U.S. Department of Energy, Washington, DC 20545, 1975.

88. BUSHONG, S. C. ET. AL. Editors, Medical Radiation Information for Litigation, DMRE 69-3, (op. cit., Reference 15), 1969.
89. WEBSTER, E. W., Patient and Personnel Dose During X-Ray Examinations, in BUSHONG, S.C., op. cit., pp. 96-129, 1969.
90. BOND, V. P. The Cancer Risk Attributable to Radiation Exposure: Some Practical Problems. Health Physics, Vol 40, 1981, pp 111-113.
91. Collins, V. P., Summary of Experience in Radiation Litigation. Published in Handbook of Management of Radiation Protection Programs, edited by: Miller, K. L. and Weidner, W. A., CRC Press, Inc., Boca Raton, FL, 1986, pp 1-27.
92. Dorothy Roberts vs. United States of America; Louise Nunamaker vs. United States of America, Civil LV 1766 RDF and 76-259 RDF, United States District Court, District of Nevada, June 14, 1984, pp 92-93.
93. BAYES, T. Essay Towards Solving a Problem in the Doctrine of Chance. Philosophical Transactions of the Royal Society, Vol 53, 1763, pp 370-418. Reprinted in Biometrika, Vol 45, 1958, pp 293-315.
94. BROWNLEE, K. A. Statistical Theory and Methodology in Science and Engineering. John Wiley and Sons, Inc., New York, 1965, pp 18-20.
95. Testimony of Dr. Allen Brodsky in the case of Dennis vs. General Electric Company, U.S. District Court. District of Maryland, Baltimore, January 11, 1984.
96. UNITED NATIONS SCIENTIFIC COMMITTEE ON THE EFFECTS OF ATOMIC RADIATION (UNSCEAR). Ionizing Radiation: Sources and Effects. 1982 Report to the General Assembly, with annexes. United Nations, New York, 1982.
97. ZEIGHAMI, E. A., WALSH, P. J., MORRIS, M. D., and JONES, T. D. Epidemiologic Measures of Risk As a Basis for Legal Compensation. In Epidemiology Applied to Health Physics. Proceedings of the Sixteenth Midyear Topical Meeting of the Health Physics Society. CONF-830101, available from National Technical Information Service, U.S. Department of Commerce, 5285 Port Royal Road, Springfield, VA 22161, 1983, pp 346-356.
98. TINSLEY, T. A. Case Histories of Compensation for Radiation Employees. In BUSHONG ET AL., op. cit., p. 200.
99. O'TOOLE, T. J. Problems of Legal Proof in Radiation Cases. In BUSHONG ET AL., op. cit., p. 263.
100. HUARD, L. A. Statutory Control of Radiation Records. In BUSHONG ET AL., op. cit., pp 349-361.
101. KRONZER, W. J. Letter in BUSHONG ET AL., op. cit., p. 374.
102. CHAVEZ, J. Remarks at the 30th Annual Conference on Bioassay, Analytical, and Environmental Chemistry. J. Chavez, Department of Energy, Albuquerque, NM, 1984.

103. FORBES, J. Remarks at the 30th Annual Conference of Bioassay ... ibid., American Nuclear Insurers, The Exchange, Suite 245, 270 Farmington Avenue, Farmington, Conn. 06032, 1984.
104. JOSE, D. Remarks at the 30th Annual Conference ... ibid., Department of Justice, Washington, DC 20530, 1984.
105. PRESTON, R. Remarks at a meeting of the Baltimore-Washington Chapter of the Health Physics Society, October 25, 1984. Science Advisor to the Senate Committee on Labor and Human Resources, U.S. Senate, Washington, DC.
106. NATIONAL COUNCIL ON RADIATION PROTECTION AND MEASUREMENTS. Management of Persons Accidentally Contaminated With Radionuclides. NCRP Report No. 65, Washington, DC: NCRP, 7910 Woodmont Avenue, Washington, DC 20014, 1980.
107. INTERNATIONAL ATOMIC ENERGY AGENCY. Manual on Early Medical Treatment of Possible Radiation Injury. Safety Series No. 47. Vienna, IAEA, 1978.
108. ROSEN, J. C., GUR, D., PAN, S. F., WALD, N., and BRODSKY, A. Long-Term Removal of ^{241}Am Using Ca-DTPA. Health Physics, Vol 39, pp 601-609, 1980.
109. NORWOOD, W. D. Therapeutic Removal of Plutonium in Humans. Health Physics, Vol 8, pp 747-750, 1962.
110. NORWOOD, W. D. J. Occupational Med., Vol 4, p 130, 1962.
111. HEID, K. R. Responses to Suspected or Known Intakes of Plutonium, in Internal Radiation Dosimetry, op. cit., in press, 1985.
112. BRODSKY, A. and HORM, I. F. Removal of ^{241}Am from Humans with DTPA. Progress Report, University of Pittsburgh, Department of Radiation Health, submitted to Department of Health, Education and Welfare, Research Grant 2R01 Ec 00122-03, 22 March 1971.
113. FASISKA, B. C., BOHNING, D. E., BRODSKY, A., and HORM, J. Urinary Excretion of ^{241}Am Under DTPA Therapy. Health Physics, Vol 21, p 523, 1971.
114. HORM, I. F. Techniques for Separating Americium from Urine After DTPA Therapy. Health Physics, Vol 21, p 41, 1971.
115. MANCUSO, T. F., SANDERS, B. S., and BRODSKY, A. "Feasibility Study of the Correlation of Lifetime Health and Mortality Experience of AEC and AEC Contractor Employees with Occupational Radiation Exposure," NYO-3394-2, 1965, p. C-5.
116. WALD, N., SPRITZER, A. A., and BRODSKY, A. "A Pilot Study of Occupational Exposure to Ionizing Radiation," Final Report, PHS Project 914.01 of the Division of Radiological Health, Contract No. SAPH 76320, Department of Occupational Health, Graduate School of Public Health, University of Pittsburgh, Pittsburgh, PA 15213, 1964.

117. MANCUSO, T. F., SANDERS, B. S., and BRODSKY, A. "Study of the Lifetime Health and Mortality Experience of Employees of AEC Contractors," NYO-3394-13, 1971.
118. STROM, D. J., BECK, W. L., STANSBURY, P. S., TANKERSLEY, W. G., and WATSON, J. E., JR. "Standard Procedures for Pooling Health Physics Data for Epidemiologic Studies," in Epidemiology Applied to Health Physics, Proceedings of the Health Physics Society Sixteenth Midyear Topical Meeting, January 9-13, 1983, CONF-830101, available from NTIS, U.S. Department of Commerce, 5285 Port Royal Road, Springfield, VA 22161, 1983.
119. JABLON, S. Characteristics of Current and Expected Dosimetry. in Prentice, R. L. and Thompson, D. J., Editors, Atomic Bomb Survivor Data: Utilization and Analysis. Philadelphia: SIAM Press, Society for Industrial and Applied Mathematics, 1984.
120. HERBERT, D. E. Model or Metaphor? More Comments on the BEIR III Report. In, Epidemiology Applied to Health Physics, op cit., 1983.
121. Reference 115, page 3-53.
122. BRODSKY, A. A Statistical Method for Testing Epidemiological Results as Applied to the Hanford Worker Population. Health Physics, Vol 36, pp 611-628, 1979.
123. BRODSKY, A. Comments on Errata Pointed Out by Frome and Khare on My Paper on Epidemiological Methods. Letter to the Editor, Health Physics, Vol 39, pp 697-699, 1980.
124. GILBERT, E. S. The Effects of Random Dosimetry Errors and the Use of Data on Acute Symptoms for Dosimetry Evaluation, in Prentice, R. L. and Thompson, D. J., op cit., 1984.
125. HEITBRINK, W. A. and CROUSE, W. E. Application of Industrial Hygiene Air Sampling Data to the Evaluation of Controls for Air Contaminants. Am. Ind. Hyg. Assoc. J., Vol 45, No. 11, pp. 773-777, 1984.
126. BEYER, W. H., Editor, Handbook of Mathematical Sciences, 5th Edition, CRC Press, Boca Raton, FL, 1978, pp. 787-788.
127. BISHOP, C. T., CASELLA, V. R., GLOSBY, A. A. and HAHN, P. B., Radiometric Method for the Determination of Uranium in Water: Single-Laboratory Evaluation and Interlaboratory Collaborative Study. EPA-600/7-79-093, prepared by the Environmental Assessment and Planning Section, Mound Facility, Miamisburg, Ohio 45342, for the Environmental Protection Agency, April 1979, p. 26 (available from the National Technical Information Service, Springfield, VA 22161).
128. AMERICAN NATIONAL STANDARDS INSTITUTE (ANSI). 1983. Criteria for Testing Personnel Dosimetry Performance, ANSI N13.11, 1983.

129. ROBERSON, P. L., HADLEY, R. T., THORSON, M. R. Analysis of ANSI N13.11: The Performance Algorithm. PNL-SA-10119, Battelle Pacific Northwest Laboratories, P. O. Box 999, Richland, WA 99352, 1982.
130. PLATO, P. and HUDSON, C. G., Performance Testing of Personnel Dosimetry Services. Published in Heaton, H. T. II, Editor, op. cit., 1982, pp 145-148.
131. MULLER, J. W. Some Second Thoughts on Error Statements. Nuclear Instruments and Methods. Vo. 163, pp. 241-251, 1979.
132. HEYDORN, K. and NORGARD, K. Analysis of Precision of Activation-Analysis methods. Talanta, Vol 20, pp. 835-842, 1973.
133. KORTE, N. and HOLLENBACH, M. Quality Assurance for Complex Chemical Analysis. American Laboratory, September 1985, pp. 152-155.

- Instruments and Methods. Vol. 103, pp. 241-251, 1973.
132. HEYDORN, K. and NORGARD, K. Analysis of Precision of Activation-Analysis methods. Talanta, Vol 20, pp. 835-842, 1973.
 133. KORTE, N. and HOLLENBACH, M. Quality Assurance for Complex Chemical Analysis. American Laboratory, September 1985, pp. 152-155.



NTIS does not permit return of items for credit or refund. A replacement will be provided if an error is made in filling your order, if the item was received in damaged condition, or if the item is defective.

technical reports.

For economy and efficiency, NTIS does not maintain stock of its vast collection of technical reports. Rather, most documents are printed for each order. Your copy is the best possible reproduction available from our master archive. If you have any questions concerning this document or any order you placed with NTIS, please call our Customer Services Department at (703)487-4660.

Always think of NTIS when you want:

- Access to the technical, scientific, and engineering results generated by the ongoing multibillion dollar R&D program of the U.S. Government.
- R&D results from Japan, West Germany, Great Britain, and some 20 other countries, most of it reported in English.

NTIS also operates two centers that can provide you with valuable information:

- The Federal Computer Products Center - offers software and datafiles produced by Federal agencies.
- The Center for the Utilization of Federal Technology - gives you access to the best of Federal technologies and laboratory resources.

For more information about NTIS, send for our FREE *NTIS Products and Services Catalog* which describes how you can access this U.S. and foreign Government technology. Call (703)487-4650 or send this sheet to NTIS, U.S. Department of Commerce, Springfield, VA 22161. Ask for catalog, PR-827.

Name _____

Address _____

Telephone _____

*- Your Source to U.S. and Foreign Government
Research and Technology.*







U.S. DEPARTMENT OF COMMERCE
Technology Administration
National Technical Information Service
Springfield, VA 22161 (703) 487-4650
