HHS Publication FDA 81-8054

11th Annual NATIONAL CONFERENCE ON RADIATION CONTROL

Radiation:
Minimizing Exposure; Optimizing Use

May 6-10, 1979

Oklahoma City, Oklahoma

June 1981

Cosponsored by

CONFERENCE OF RADIATION CONTROL PROGRAM DIRECTORS, INC.

•

ENVIRONMENTAL PROTECTION AGENCY
Office of Radiation Programs

•

U.S. NUCLEAR REGULATORY COMMISSION

and

U.S. DEPARTMENT OF HEALTH AND HUMAN SERVICES
Public Health Service
Food and Drug Administration

Bureau of Radiological Health
Rockville, Maryland, 20857

REPORTING of ENVIRONMENTAL RADIATION MEASUREMENT DATA

R. Collé
Office of Radiation Measurement
Center for Radiation Research
National Bureau of Standards

The Task Force on Radiation Measurements of the Conference of Radiation Control Program Directors in its first report to the National Bureau of Standards (NBS) in February 1977 recommended that:

> "NBS in cooperation with EPA should be designated as the Federal agency responsible for coordinating efforts among the Federal and State agencies to immediately develop a uniform data reporting system so that present environmental data which are being generated throughout the country can be utilized and evaluated in terms of possible population exposure."

At approximately the same time, a Health Physics Society ad hoc committee was organized for the purpose of upgrading environmental radiation data. One of the 9 project objectives identified by that committee had a scope nearly equivalent to the Task Force recommendation. Specifically, the HPS committee identified the need to "establish more standardization in data reporting conventions including the definitions of detection limits and the treatment of results at or below these limits." As a result, NBS agreed to coordinate and implement the project under the auspices of the HPS committee and an appropriate subcommittee was formed to serve as a working group (1). This subcommittee received very broad national participation by representatives of private industries, government contract laboratories, State governments and Federal agencies including DOE, EPA, NBS, and NRC.

A draft of a document, A Guide and Recommendations for Reporting of Environmental Radiation Measurements Data (2), prepared by the subcommittee has been completed and distributed to a wider audience for review and comment. This guide and its detailed recommendations are intended to be useful and specific. It is not intended to be a general or exhaustive theoretical treatment of measurement results, but rather to serve as a practical guide to treating and reporting environmental radiation measurements data. A tutorial philosophy and approach was adopted because it was envisaged that many of the intended users would be unfamiliar with some of the relatively complex concepts and subjects covered in the Guide. Emphasis throughout has been placed upon obtaining solutions and not just reidentifying the problems associated with reporting data. Present data reporting practices are summarized and evaluated, and their deficiencies are examined. In those cases where current practices are unjustified, the reason for alternative approaches are carefully explained. Nearly 30 specific recommendations for a uniform method of data reporting are presented and justified. From the onset, the subcommittee recognized that its recommendations will not be adopted unless they are strongly justified. Additionally, the Guide contains considerable supporting materials so that it may be used without recourse to other reference sources and extensive examples to demonstrate the practical use of the recommended methods.

The complete Guide will be published as an NBS Special Publication (2) and will be available in the Fall of 1980. A condensed version will be included, along with eight other reports, in a publication of the Health Physics Society that will present the combined output of the committee on upgrading environmental radiation data (3). Highlights from the Guide which are still provisional follow.

exposure, absorbed dose, and dose equivalent. The status of each of these, in terms of both the old units and new SI units, is considered with the Guide.

The Guide recommends that data reports should convert to the SI base and derived units for radiation quantities as soon as it is technically and economically feasible. In view of existing governmental regulations and data-reporting requirements, it is, at present, neither feasible nor economically justifiable to completely and immediately change over from old to new units. This change will require familiarization and a transition period. Efforts to plan and implement a safe and orderly transition to SI units should be initiated by concerned laboratories, independent organizations and governmental agencies to aid and expedite the transition.

A consistent set of units for various environmental radiation quantities (for both the old and new SI units) is contained with the Guide and should be employed for data reporting. In addition to presenting these recommended units, the Guide also contains (1) a summary of dimensioning conventions including the treatment of compound units, multiples and submultiples of units, and special units, (2) guidelines for expressing results in exponential notation, (3) some considerations for the transition to the use of SI units, and (4) a discussion of some unsatisfactory current practices such as the extensive reporting of "gross activity" measurement results in units of activity.

II. SIGNIFICANT FIGURES

The second requirement (outlined in the Overview) for reporting measurement data was that the reported value must be "expressed in an appropriate number of significant figures." Fortunately, this is nearly universally recognized, and is seldom a problem.

The "number of significant figures" refers to the number of numerical digits that is used to express the value. It is obvious that this number must be reasonable, and must not mislead or imply fictitious accuracy in the reported value. An "appropriate" number of digits (or significant figures) is that which is warranted by the accuracy of the reported value. That is, the appropriate number of digits to be retained in the reported value depends on the magnitude of the total uncertainty (see Part III) attached to this value.

In general, environmental radiation measurements data seldom justify more than two or three significant figures for the value and one or two significant figures for the uncertainty. More significant figures should only be reported after careful consideration and a decision that the extra figures are indeed reportable. It is recommended that the uncertainty should be reported to no more than two significant figures, and the value itself should be stated to the last place affected by the qualification given by the uncertainty term. The value and its uncertainty must be in decimal agreement. Some consistent, well-developed rules for rounding reported values and their uncertainties to an appropriate number of significant figures are provided in the Guide.

These views are currently incorporated into most data reporting practices and regulatory requirements.

III. TREATMENT OF UNCERTAINTY STATEMENTS

Three of the more common abuses concerning the reporting of uncertainties for environmental radiation measurements data were summarized in the Overview. These abuses are addressed in detail within the Guide, and will not be considered here. It may be useful to state, however, the necessary conditions for avoiding their failings:

A REPORTED VALUE MUST INCLUDE AN ASSESSMENT OF ITS UNCERTAINTY.

sive independent measurements. Precision then is a measure of the closeness together. Random uncertainty components can be considered to be those sources of inaccuracy which can be and are assessed and propagated by statistical methods. Sample statistics, such as the standard deviation s_x, which are computed entirely from the measurement data, are used to estimate the population parameters such as σ_x. That is,

$$\text{SAMPLE STATISTIC} \rightarrow \text{POPULATION PARAMETER}$$

$$(\text{e.g., } s_x) \qquad\qquad (\sigma_x)$$

7. A bias is a deviation from τ, which is always of the same magnitude and direction. It cannot be estimated or calculated solely from a given set of replicate measurements since each and every measurement is affected by the systematic bias in the same way. For a set of replicate measurements with a limiting mean of μ_x associated with the measurement of the particular quantity by the given measurement process, a bias is the difference between the limiting mean (μ_x) and the true value of the quantity,

$$\text{BIAS} = \mu_x - \tau.$$

8. For completeness, the systematic uncertainty components (δ) can be considered to be the residual set of the conceivable sources of inaccuracy which are biased and those which may be due to random causes or stochastic processes but cannot or are not assessed by statistical methods. Therefore,

$$\begin{matrix}\text{SYSTEMATIC}\\ \text{UNCERTAINTY}\end{matrix} \rightarrow \left(\text{BIAS} \quad + \quad \begin{matrix}\text{UNEVALUATED}\\ \text{RANDOM}\\ \text{COMPONENTS}\end{matrix}\right)$$

9. The total uncertainty used as an estimate of the inaccuracy is then just some combination of the random and systematic uncertainties, i.e., should be:

N.B. symbols missing

$$\begin{matrix}\text{TOTAL}\\ \text{UNCERTAINTY}\\ (\mu_x)\end{matrix} \cancel{=} f\left\{\begin{matrix}\text{RANDOM}\\ \text{UNCERTAINTIES}\\ (s_x)\end{matrix}, \begin{matrix}\text{SYSTEMATIC}\\ \text{UNCERTAINTIES}\\ (\delta)\end{matrix}\right\} \qquad \mu_x = f\{s_x, \delta\}$$

These concepts are summaried in Figure 1. The total uncertainty for a measurement result is ultimately used to estimate the accuracy of the result. It is obtained by combining the random and systematic components. The random uncertainty is a measure of precision in which calculated sample statistics are used to estimate population parameters. The systematic uncertainty includes sources of inaccuracy in addition to biases in order to obtain a complete accounting of all sources of inaccuracy in the total uncertainty. The very familiar "bull's-eye" example shown in Figure 2 should help illustrate the distinction between precision and systematic biases, and their relation to accuracy. The bull's eye of the target corresponds to the true value, and the six shots represent individual measurement results. The figure illustrates the concept of an inaccurate measurement due to imprecision, a precise but inaccurate measurement, and an accurate measurement. Inasmuch as accuracy requires precision, there is no such case as an accurate but imprecise measurement. It should be mentioned that a third category of conceivable sources of inaccuracy which has not been mentioned previously is that termed "blunders." These are the outright mistakes in the measurement process. If one or more blunders are known to exist, then the value should be discarded, or a correction should be applied to the measurement. One must recognize that blunders do and will occur. The Guide recommends establishing an internal system of independent checks on each stage of the measurement and data reduction system to detect and avoid blunders.

A large part of the Guide is devoted to the process of making a complete uncertainty assessment of all of the conceivable sources of inaccuracy, and the mechanics of treating the individual components from their initial assessment through their propagation. Figure 3 attempts to illustrate the concept of making this complete assessment, and the division of the sources of inaccuracy between the random and systematic components. Every conceivable source of inaccuracy should be classified into one of the two categories depending on how the uncertainty is estimated. Until combined for the measurement result, the individual uncertainty components in both categories should be retained separately.

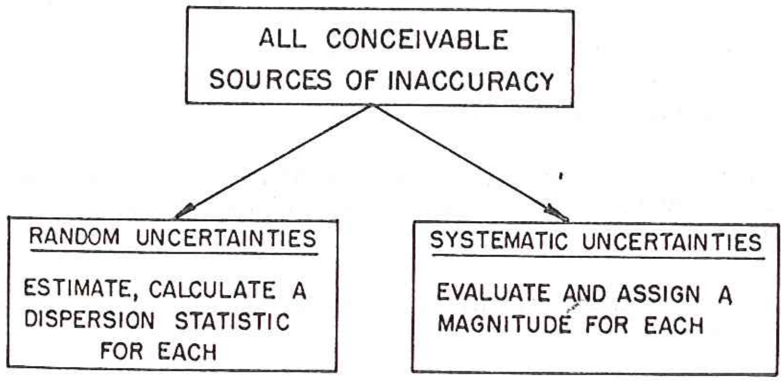

Figure 3

Each random uncertainty component should be estimated in terms of its standard deviation, or the standard error of the mean for independent multiple measurements. These components are estimated by a statistical analysis of replicate measurements. The random "counting error" contribution may be estimated from single measurements by making the familiar Poisson assumption. Although the use of the term "counting error" will probably continue to prevail, the term is a gross misnomer. It takes into consideration only the random scatter about the mean from the radioactive decay process itself. To presume that this is the only source of random fluctuation in the overall measurement process is fatuous. Unfortunately, Federal guidance, as it pertains to reporting environmental radiation data, does not suggest otherwise. Other sources may be random timing uncertainties, random variations in the sample preparation, positioning of the sample at the detector, and so forth. The list is nearly endless. Many of these are addressed in a detailed appendix to the Guide.

The only way to realistically assess the overall random uncertainty is to make replicate determinations and calculate the standard deviation of the mean by the usual statistical methods. For reported values derived from several component quantities, the individual uncertainty estimates (obtained from replicate measurements for each independent variable) must be propagated to obtain the overall standard deviation of the final value. It is recognized that replicate measurements of environmental samples are uncommon and may not be feasible because of time and cost constraints. Yet an uncertainty assessment based on some procedure beyond calculating the square root of the total number of counts is clearly needed. The Guide recommends that the total random uncertainty be obtained by propagating the individual variances for each random uncertainty component. This would include not only the Poisson counting error and random uncertainty derived from the particular measurement under concern, but also the random components from corrections, constants, calibration factors, and any other measurements that make up the final result. These contributions may be combined by summing in quadrature or by the use of other propagation of error formulae given in the Guide. It should be emphasized that one benefit of making a complete assessment is that the process will require the evaluation and statistical control of many previously unevaluated measurement parameters. This will ultimately improve the quality of measurements and aid quality control within laboratories.

are interrelated. Underlying this entire discussion is the fundamental consideration that the way uncertainties are reported is dependent on their intended or ultimate use. A number of common practices and methods of combining the random and systematic uncertainty components were considered and reviewed in the Guide. It became apparent that there is no one clearly superior method out of the range of approaches to reporting an overall uncertainty. Yet, the diversity of current practices and methods either recommended or required by governmental bodies, or used by those who report data make it difficult if not impossible to combine or compare these data in a manner which would be useful for determining trends or possible necessary actions. There is, therefore, a great need for uniformity in the method used to report the uncertainties of environmental radiation data. In hope of achieving greater uniformity, a method, often termed the PTB approach (4), is recommended. In its simplest form, the random and systematic uncertainties should be combined to form an overall uncertainty in the result x by

$$u_x = \sqrt{s_x^2 + 1/3 \sum_{j=1}^{q} \delta_j^2}$$

where s_x is the standard deviation corresponding to the overall random uncertainty and δ_j is the magnitude of the estimated systematic uncertainty for each of the q systematic uncertainty components. Additional details and other special cases are given in the Guide. The recommended method requires that each reported measurement result include the value, the total random uncertainty expressed as the standard deviation, and a combined overall uncertainty. Examples illustrating the method are given in an appendix to the Guide. This recommended approach may not be the best for all purposes. It is intended to be useful and practical without imposing unnecessary burdens on the time, money, and personnel resources of laboratories. Adoption of this method may eventually demonstrate its most serious shortcomings and ultimately lead to better methods.

IV. DETECTION LIMITS

A myriad of vastly different expressions and definitions of "detection limits" are frequently encountered. Their meanings are often ambiguous, inconsistent, and incorrectly interpreted. Some of these include "detection sensitivity," "minimum detectable activity (or level)," "lower limit of detection," and "background equivalent activity." Some of the problems and inconsistencies in their use are addressed in the Guide.

Pragmatically, a detection limit is useful as a criterion for experiment design, comparison, and optimization purposes, such as in selecting among alternative measurement procedures. Additionally, detection limits may serve as guides which are set by regulatory bodies for establishing minimum acceptable detection capabilities for a given type of analysis. For the purposes of the Guide, the intent of detection limit calculations is to satisfy both of these uses. It must be emphasized that any calculation of a detection limit is at best only an estimate. Their use is limited to that of serving as guideposts only, and not as absolute levels of activity that can or cannot be detected by a counting system.

Much of the existing confusion with detection limits for environmental ionizing radiation measurements arises from not only the large number of different expressions that are in use, but also from incorrect interpretations and misapplications of some of the original definitions. In order to satisfy both of the above mentioned purposes, two distinctly different concepts are required.

The first is an estimated detection limit that is related to the characteristics of the counting instrument. It is not dependent on other factors involved in the measurement method or on the sample characteristics. It is a lower limit in the true sense of the word. Because of its current wide usage, the recommended term is the ESTIMATED LOWER LIMIT OF DETECTION (LLD).

As stated earlier, the LLD is an a priori estimate dependent on only the instrument background and detection efficiency, and the MDC is an a priori estimate for a given type of analysis or measurement process under specified typical conditions. They are not a posteriori decision limits for every measurement. Therefore, they need not, and should not be calculated for each individual measurement. The practice of comparing a computed LLD or MDC for each measurement against the measurement result should be avoided. This has sometimes been employed to determine the "significance" of the result for reporting purposes.

If results below a particular computed LLD or MDC are rejected or excluded from data reports, serious errors and distortion in long-term trends could result. Consider the hypothetical situation illustrated in Figure 4. The background is represented by the probability distribution with mean μ_b and standard deviation σ_b. Similarly, a gross count is given by the distribution with μ_g and standard deviation σ_g. Thus, the net count, obtained by subtracting the background from the gross count, is the probability distribution with mean $\mu_n = (\mu_g - \mu_b)$ and variance $\sigma_n^2 = (\sigma_b^2 + \sigma_b^2)$. This net count distribution can be compared, as shown, to the estimated LLD calculated with the standard deviation of the background (σ_b). In this case, the true or limiting mean of the net count, and a substantial fraction of the distribution is below the computed LLD. A resultant net count (N_n) from a typical individual paired observation of background and gross counts (N_b, N_g), which are both well within $\pm\sigma$, is illustrated. If this value and many others less than the LLD are not reported then the distribution of the reported values would grossly distort the true situation given by μ_n and σ_n for the population. This point is demonstrated further in the data of Table 1. Twenty-five net count (N_n) measurement results from a population with a mean of 20 were obtained from paired observations of background (N_b) and gross counts (N_g). The average background of 16 counts was used to estimate the LLD, which has a value of 21.3 or 18.6 counts (depending on the particular equation used to calculate the LLD). The last two columns in the table contain the tabulated results when values less than 21.3 and 18.6 counts are reported as "less than the LLD" (<LLD). In the first case, only 8 of 25

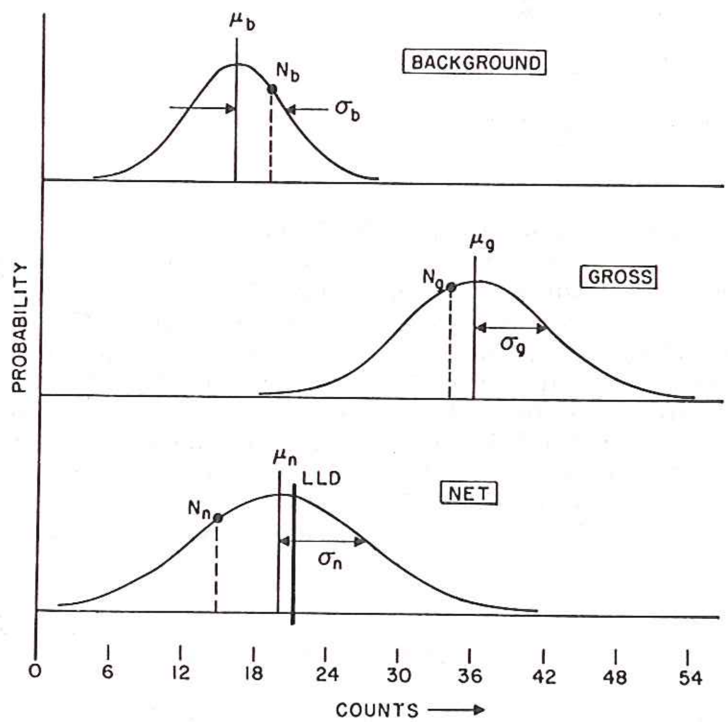

Figure 4

it does not reflect realistic detection capabilities. It is sometimes argued that the LLD must consider that any net count results from contributions from both strontium-89 and strontium-90. This argument is another misapplication of the LLD concept. If it is accepted, then the LLD would no longer be a limit and would be dependent on something other than the instrument background and efficiency characteristics.

The MDC, however, is intended to serve as a practical level that is achievable under a given set of specified measurement conditions, and must consider the effect of multi-components in the sample. One is therefore confronted with a paradox. The MDC should be an a priori estimate, but its determination requires a posteriori knowledge of the amounts of activity in the sample. This dichotomy can be overcome by calculating the MDC for typical sample conditions. This is subject to criticism that some samples may vary substantially from the typical, and that the MDC may then be in considerable error. Without denying this argument, there are two reasons why the effects from such errors are unimportant. First, the MDC is only meant to be an **estimate**; and second, all measurement results should be reported without a comparison to the MDC. As a result, there is no great need for knowing some absolute detection level for every measurement.

The determination of a MDC for activity measurements by gamma-ray spectrometry is a good example of the above described situation. It is complicated in that the background count rate for the same radionuclide may vary from sample to sample. This is because the background is usually the sum of two separate sources of gamma radiation. First, there is the system background radiation for a blank, corresponding to the sample to be measured, as discussed previously. Second, there is Compton-scattered radiation from other higher energy gamma rays in the sample. Thus, it is readily seen that the measurement of blank backgrounds has meaning only for very specific situations; e.g., when only one radionuclide is present in the sample, or when the radionuclide of interest emits the highest energy gamma ray.

For environmental samples, one can assume that almost always two or more gamma-emitting radionuclides will be present in any given sample. Use of the uncertainty in the blank background is appropriate for the LLD, but a more realistic background must be used for calculating a MDC. This sample background will depend upon the relative amounts of other radionuclides that are present in the sample and the energy of each of their gamma rays. Therefore, to satisfy the condition that the MDC be an a priori estimate, it is recommended that it be calculated for each radionuclide from typical spectrum backgrounds for each type of sample.

In summary, the practical significance of the estimated LLD and MDC are only to serve as guideposts or criteria for experiment design, comparison, and optimization purposes, and to serve, for regulatory purposes, as approximate guidelines of minimally acceptable levels that can be practically achieved. As such, all measurement results should be reported directly as obtained, and the estimated LLD or MDC should not be employed to exclude some results from reports. Similarly, the practice of calculating a new LLD or MDC value for every measurement defeats the spirit of the concepts, which is to provide a priori estimates. A posteriori calculations contain no additional information, and are neither technically nor economically justifiable.

The above interpretations place some important restrictions on the use of the LLD or MDC, particularly with respect to satisfying regulatory specifications. The most important restriction is that it is unreasonable for a regulator to establish absolute values for the MDC or LLD for various combinations of radionuclides and type of sample medium, and then expect compliance 100 percent of the time in all situations.

A further restriction arises from the fact that the MDC for a radionuclide may vary from sample to sample depending on the characteristics of the sample. For example, the MDC for strontium-90 is dependent on the amount of strontium-89 which is also present in the sample. Similarly, the MDC for a radionuclide assayed by gamma-ray spectrometry

3. Upgrading Environmental Radiation Data, Committee for Upgrading Environmental Radiation Data, Health Physics Society, to be published (1980), Chapter 6.

4. S. Wagner, PTB Mitteilungen 79, 343 (1969). An English version, "How to Treat Systematic Errors in Order to State the Uncertainty of a Measurement," is available from the Physikalisch-Technische Bundesanstalt, Braunschweig, as report number FMRB 31/69. Additional details may be found in P.J. Campion, J.E. Burns, and A. Williams, "A Code of Practice for the Detailed Statement of Accuracy," National Physical Laboratory, London (1973); and A. Williams, P.J. Campion, and J.E. Burns, "Statement of Results of Experiments and Their Accuracy," Nucl. Instr. and Meth. 112, 373-376 (1973).

5. Lloyd A. Currie, Anal. Chem. 40, 586-93 (1968).

6. B. Altshuler and B. Pasternack, Health Phys. 9, 293-8 (1963); B.S. Pasternack and N.H. Harley, Nucl. Instr. Meth. 91, 533-40 (1971).

MR. CRAIG: I am glad to see you presenting these kinds of things and for some of your cohorts in the EPA who have never heard of negative numbers, the Model States Information System which is designed to serve as a data base for the Drinking Water Act has been written in such a fashion that it will not accept negative numbers. I would like to see the NBS and Office Radiation Programs discuss this and convince them that they are going to have a long-term bias and are going to come up with concentrations that are going to give everyone trouble with some new standards.

DR. COLLE: If this is the case, then I would say that there are going to be some problems. Serving on our committee were two EPA representatives; otherwise it strikes me as being matters that are internal to EPA.

MR. WOLF: I think it is extremely important to discuss in some detail what type of environmental quality assurance programs you have, whether you are splitting samples and how many, whether it is 20 percent of the samples that one is collecting and reading. A lot of environmental data are reported without any discussion of the quality control program.

DR. COLLE: A quality control program would literally be used for the most part to evaluate many of the sources of inaccuracy. The Guide contains a detailed example of the assessment and propagation of the uncertainty components for strontium-89 and strontium-90 determinations. Nearly 20 different sources of inaccuracy are identified. Each source has to be evaluated. One of the ways to evaluate them is through various quality assurance programs as well as maintaining inhouse statistical control. Inhouse quality controls and external quality assurance programs by splitting samples and cross-checks will also help evaluate these terms, and over time will also improve the quality of measurements.