Empirically Determined Decision Levels
Development and Use in an *In Vivo* Bioassay Program

Brian J. Lawson, Michael A. Orcutt and Richard J. Winslow
Lockheed Martin Corporation
PO Box 1072, Schenectady, NY 12301

ABSTRACT

This paper discusses empirically determined decision levels in an *in vivo* bioassay program. Specifically, the paper provides an overview of ANSI N13.30 decision level concepts, the process used to determine empirical *in vivo* decision levels using the Canberra Industries ABACOS PLUS software, and the use of the decision levels in evaluating and reporting personnel results. The paper includes the development of empirically determined decision levels for uranium (U-235 and Th-234) using a Low Energy Lung Monitor (LELM), for cesium-137 using a Canberra Industries ACCUSCAN-II whole body counter, and for cobalt-60 using a Canberra ACCUSCAN-II operating in a stationary detector mode as a lung counter.

BACKGROUND

Lloyd Currie's work[1] on radioassay result reporting provides the statistical basis for ANSI N13.30 and ANSI N42.2. ANSI N13.30 **Performance Criteria for Radiobioassay** has been published for nearly ten years, having been revised numerous times. It still exists as a draft standard. ANSI N42.2 **Measurement Quality Assurance for Radioassay Laboratories** has recently (2/9/94) been issued as a final revision. Both of these documents provide important guidance to internal dosimetrists and others. It is just recently, nearly thirty years after publication, that Currie's concepts are becoming widely used by analysts.

Key to Currie's work is the decision level concept. The decision level, as defined in ANSI N42.2, is that quantity of analyte at or above which an *a priori* decision is made that a positive quantity of the analyte is present. An *a priori* decision is one made prior to the measurement (as compared to *a posteriori*, or after the measurement).

ANSI N13.30 discusses appropriate blanks for a sample, person, or phantom. An appropriate blank is, ideally, identical in physicochemically and radiologically significant ways to the sample or person of interest. But what is an appropriate blank for an *in vivo* program? And, how does a laboratory reasonably arrive at a decision level when there is variability between counting subjects? Variability can arise from diet, water supply, physiological differences, etc. How does an internal dosimetrist take all this into account when running an *in vivo* program? We propose the use of an empirically determined decision level.

RADIOLOGICAL WORKERS, NON-RADIOLOGICAL WORKERS, and EMPIRICAL DECISION LEVELS

The purpose of an *in vivo* bioassay program is to provide an assessment of whether radiological

---

[1] **Analytical Chemistry**, Volume 40, Number 3, March 1968

workers have been exposed to operational internal radioactivity, and if so, what the resulting dose is. The keyword is operational. The internal dosimetrist is not interested in non-operational radioactivity. With the analytical system sensitivities available today, we are able to measure very low levels of radioactivity, including levels that exist as naturally occurring radioactivity, and which are indistinguishable from operational radioactivity.

The expectation we have is that, overall, results of routine lung scans and whole body counts for radiological workers are no different than those of non-radiological workers. This is because the radiological engineering and work controls used for the performance of radiological work preclude the potential of internal exposure. Stating this statistically, we do not expect the distribution of baseline results for people who have never handled the nuclide of interest to be any different from the distribution of routine samples for radiological workers. This can be tested by the following equation:

$$T = \frac{(\overline{X}_1 - \overline{X}_2) - (\mu_1 - \mu_2)}{\sqrt{\frac{S_1^2}{n_1} + \frac{S_2^2}{n_2}}}$$

where:

T is the calculated test statistic,
$\overline{X}_1$ is the mean of data available for distribution 1,
$\overline{X}_2$ is the mean of data available for distribution 2,
μ_1 is the true mean of distribution 1,
μ_2 is the true mean of distribution 2,
S_1^2 is the variance of distribution 1,
S_2^2 is the variance of distribution 2,
n_1 is the number of results in distribution 1,and
n_2 is the number of results in distribution 2.

The null hypothesis is that $\mu_1 - \mu_2 = 0$; that is, there is no difference in the two means. To determine the test statistic (t) determine $t_{\upsilon,1-\alpha/2}$, (using a Student t Distribution table) where $\alpha = .05$ and υ is determined by:

$$\upsilon = \frac{(a_1 + a_2)^2}{\frac{a_1^2}{(n_1-1)} + \frac{a_2^2}{(n_2-1)}}$$

where, $a_1 = S_1^2/n_1$ and $a_2 = S_2^2/n_2$. If $T < t_{\upsilon,1-\alpha/2}$, there is no statistically significant difference in the means of the two distributions, and our null hypothesis is upheld. Alternatively, if $T > t_{\upsilon,1-\alpha/2}$, there is a statistically significant difference in the means of the two distributions.

This expectation of the equality of these two sample distributions is the basis for the development of what we call the empirical decision level.

Section A.7.3 of ANSI N42.2, **Interpretation of Individual Measurement Results**, states:

> "For the purpose of having a laboratory interpret whether an individual sample measurement is different from its representative appropriate blank, it is recommended that the laboratory compare the net count or count rate of the measurement with a decision level calculated using the sample specific "appropriate blank". The "appropriate blank" should include measurement interferences from impurities that are not typically known *a priorily* or included in the *a priori* decision limit. This "true" decision level is different from the nominal *a priori* decision level in that it truly represents the appropriate blank at the time of measurement. For some measurement processes, the determination of the "true" appropriate blank for each sample may be impractical."

We consider that for **radiological worker** *in vivo* bioassay, the distribution of **non-radiological** worker results for the corresponding analytical process can be treated as the "appropriate blank". The decision level for radiological worker *in vivo* bioassay can be estimated by counting a population of non-radiological workers. The non-radiological worker results contain the interferences from impurities (and naturally occurring levels of radioactivity) that are typically present in operational lung scans and whole body counts.

If, as we stated earlier, Currie considers the decision level as that quantity of analyte at or above which an *a priori* decision is made that a positive quantity of the analyte is present, and we are willing to accept a 5% chance of a Type I (false positive) error, then the "true" decision level can be estimated by the 95th percentile of the distribution of results for non-radiological workers. By setting the empirical decision level at the 95th percentile of the distribution of results for non-radiological workers, we accept a false positive rate of approximately 5%. It is the 95th percentile of the distribution of results for non-radiological workers that is used to check individual results for radiological workers. Results below the decision level indicate that the subject is indistinguishable from the unexposed population from a bioassay standpoint, and followup is therefore not warranted.

As recommended in ANSI N13.30, Appendix A, equation A.9, the decision level can be calculated by:

$$L_c = 2.33\, s_{b,}$$

where s_b = standard deviation of the blank counts.

In addition, the decision level can be estimated from the count result of the 95th percentile result from the counts of the unexposed population. Comparing the calculated decision level and the 95th percentile result serves as a good crosscheck of the population selected (i.e., it is large enough, follows a normal distribution, and does not have significant anomalies).

A critical point to make regarding empirical decision levels is that an empirical decision level is specific to a nuclide, an assay procedure, and the performing laboratory (i.e. counting equipment used). Any change to the analytical process could be reason to reestablish the decision level.

METHODOLOGY FOR DETERMINING EMPIRICAL DECISION LEVEL(EL_c)

Analytical Systems - Hardware

The nuclides of interest which empirical decision levels were established are uranium-235 (^{235}U), uranium-238 (^{238}U based on the thorium-234 (^{234}Th) daughter), cesium-137 (^{137}Cs) and cobalt-60 (^{60}Co). Two systems were utilized in this study, both built by Canberra Industries. The following is a description of the two systems:

Low Energy Lung Monitor (LELM)

The LELM is a state-of-the-art *in vivo* monitoring system designed primarily for the detection of ^{235}U and selected transuranics. The system is comprised of an iron room with 6.4 inch thick, low radioactivity iron walls, ceiling, and floor and a 0.5 inch thick layer of low radioactivity lead on all inner surfaces.

Two, four-detector arrays (8 detectors in all) of hyperpure, low energy germanium detectors (LeGe) are used to perform lung monitoring. The detectors have an automatic liquid nitrogen (LN_2) fill system and detector protection circuits to ensure that the LeGes operate at the required LN_2 temperatures. This LN_2 system is rendered inoperable during subject counting by a pressure sensitive interlock switch, installed in the subject chair for the subject's safety.

Each detector has an active diameter of 50.9 mm, which corresponds to a total active area of slightly greater than 2,000 mm^2 for all eight detectors. The thickness of each detector is 15 mm. The distance from the inner surface of the 0.5 mm beryllium window to the detector is 4 mm. This distance allows for slight flexing of the window without detector damage. A typical resolution achieved by these detectors at 5.9 keV is 350 eV full width at one-half maximum (FWHM) and 600 eV FWHM at 122 keV.

High Energy Lung Monitor(HELM) and Whole Body Counter (WBC)

The Accuscan II is a state-of-the-art *in vivo* monitoring system designed primarily for the detection of higher energy (e.g. ^{60}Co, ^{137}Cs) activated corrosion products and mixed fission products. The system is comprised of a steel tub with 4 inch thick, low radioactivity steel walls, ceiling, and floor.

Two closed-end coaxial hyperpure germanium detectors are used to perform whole body and lung monitoring for higher-energy gamma and X-ray emitting radionuclides. The relative efficiency (to sodium iodide) of these detectors is 25%. A typical resolution achieved by this system at 122 keV is 0.9 keV FWHM and 1.9 keV FWHM at 1332 keV. With the detectors stationary over the lungs, the system is highly effective as a HELM. With the detectors in motion from head to toe, it is an effective whole body counter (WBC).

Computing Equipment

A Digital Equipment Corporation (DEC) microVAX 3400 computer, an X-terminal, and a DEC VAXstation 4000 are used to process *in vivo* monitoring files, calibration spectra and data, quality assurance data, and results. An Okidata OL830 Plus is used to print out ABACOS PLUS reports and Spectra. The microVAX 3400, VAXstation 4000, LELM detectors, ACCUSCAN II detectors, ACCUSCAN II detector motion control, and the X-terminal each are a node on a 50 ohm Thinwire Ethernet network.

Analytical Systems - Software

The analysis software utilized on both systems is Canberra Industries ABACOS PLUS, a state-of-the-art *in vivo* counting software application. It provides the software functions needed to perform *in vivo* measurements of nuclide activity and calculate corresponding internal doses, if required. It provides menu-format options to create and calibrate counting systems using various combinations of hardware. The program has been customized for decision level reporting.

Gamma M is a peak search algorithm within ABACOS PLUS that allows the user to define certain parameter values. The ability to adjust sensitivity parameters relating to peak identification is the key to determining and utilizing empirically determined L_c values using the Canberra ABACOS PLUS *in vivo* counting software. The user definable parameters used are "Reject MDA sigma" and "Reject MDA constant". These parameters are used in locating potential peaks during the library-driven peak search routine. These parameters specify how large the net peak area must be, relative to the standard deviation of the underlying background continuum, to be retained and reported as statistically significant (and hence, be used in calculating an activity). Peaks will be identified if:

$$\text{Net Peak Area} > ((\text{Reject MDA sigma}) * (S_b)) + \text{Reject MDA constant.}$$

As the two parameters are decreased, the net peak area becomes larger relative to the screening criteria value, and the system becomes more "sensitive" to identifying peaks.

Process

The methodology for determining an empirical decision level for *in vivo* counting systems is:

1.) Personnel at your facility who have never been operationally exposed to the nuclide of interest are identified.

2.) A statistically meaningful number of counts (minimum of 40-50) are performed using the personnel identified above as subjects.

3.) When counts are performed for this study, they should be performed using the same analytical process that your operational subjects will be subject to.

4.) Results of the decision level samples are reported as counts in the region of interest.

5.) Once the population of non-radiological worker data has been collected, the data should be

ranked in order of results. Statistical outliers, if any, should be discarded using standard statistical methodologies.

6.) The empirical decision level is determined by calculating the standard deviation of the non-radiological worker population of results, and multiplying this by 2.33.

DATA - CESIUM - 137

One hundred counts were performed over a period of about six months. The one hundred spectra were analyzed using Canberra Industries ABACOS PLUS with Reject MDA Sigma and Reject MDA Constant set to the vendor recommended values for MDA reporting. The total number of counts within the ^{137}Cs region of interest were extracted from the spectrum and ranked in ascending order (see TABLE 1).

Table 2 provides summary data and calculational results for standard deviation and the empirical decision level (EL_c).

Table 1: Cesium -137 Whole Body Count - CANBERRA ACCUSCAN II - ABACOS PLUS

#	Counts	#	Counts	#	Counts	#	Counts
1	0.000000	26	1.939940	51	3.155030	76	4.000000
2	0.000000	27	2.000000	52	3.155030	77	4.000000
3	0.000000	28	2.000000	53	3.155030	78	4.000000
4	0.000000	29	2.000000	54	3.155030	79	4.000000
5	0.000000	30	2.000000	55	3.155030	80	4.155030
6	0.155029	31	2.000000	56	3.155030	81	4.155030
7	0.155029	32	2.000000	57	3.155030	82	4.155030
8	0.155029	33	2.155030	58	3.629880	83	4.310060
9	0.155029	34	2.155030	59	3.629880	84	4.310060
10	0.629883	35	2.155030	60	3.629880	85	4.310060
11	0.784912	36	2.155030	61	3.629880	86	4.629880
12	1.000000	37	2.155030	62	3.629880	87	4.629880
13	1.000000	38	2.310060	63	3.629880	88	4.784910
14	1.000000	39	2.629880	64	3.784910	89	4.784910
15	1.000000	40	2.629880	65	4.000000	90	4.784910
16	1.000000	41	2.629880	66	4.000000	91	5.629880
17	1.000000	42	2.629880	67	4.000000	92	5.629880
18	1.000000	43	2.629880	68	4.000000	93	5.629880
19	1.155030	44	2.629880	69	4.155030	94	5.629880
20	1.310060	45	2.784910	70	4.155030	95	5.629880
21	1.310060	46	2.784910	71	4.155030	96	6.000000
22	1.629880	47	3.000000	72	4.310060	97	6.629880
23	1.629880	48	3.000000	73	4.310060	98	7.000000
24	1.784910	49	3.000000	74	4.310060	99	9.629880
25	1.784910	50	3.000000	75	4.629880	100	9.939940

Note that fractional counts are the result of the software determining the number of counts in a region of interest based on the energy calibration of a nuclide and not the associated channel number.

Table 2: Cs-137 Summary

Entity	Value (*counts)
Cs-137 Avg.	3.11*
S_b	1.98
Cs-137 Min	0*
Cs-137 Max	9.94*
$EL_c = 5^*$	

Using the Currie equation to calculate the decision level counts in the region of interest yields (from Table 2) 2.33 * 1.98 = ~ 5 counts. The 95th percentile is about 6 counts in the region of interest, using the empirical data directly (see Table 1 shaded value). These results are in adequate agreement. A decision level of 5 counts, by calculation, was anticipated to yield a false positive rate of about 10% based on a review of the empirical data. Therefore, 6 counts was selected as the decision level.

The empirical decision level value of 6 counts was then used to calculate a corresponding activity. The activity corresponding to 6 counts in the region of interest for a 10 minute whole body count is 2.7 nanocuries (nCi). This calculated activity becomes the Empirical Decision Level (EL_c) activity for the nuclide of interest. The EL_c is the value at or above which a decision is made that counts in the region of interest are present above the 95th percentile for a non-occupationally exposed population, and a recount will be performed.

The final step is to adjust the Reject MDA Sigma and Reject MDA Constant user definable parameters in the Gamma M peak search routine so that when counts in the region of interest are greater than or equal to the EL_c counts, a peak is "found", and the analysis reports a "positive" (i.e., greater than decision level) value. Selection of Reject MDA Sigma and Reject MDA Constant is an empirical process and can not be demonstrated mathematically. Various values are input until the top 5 spectra are reported as "positive".

IMPLEMENTATION OF DECISION LEVEL REPORTING

ANSI N.42.2 provides recommendations on the interpretation of radioassay results. Specifically, as mentioned in section A.7.3, the laboratory should compare the sample count or count rate to the decision level count or count rate using an appropriate blank. The empirical decision level, determined as above, becomes the screening level for reporting results as "positive" or negative. Again it is emphasized that a ~5% false positive rate is built into the process.

Although ANSI N42.2 is moot regarding *in vivo* analysis specifically, we consider these same

principles can be applied to *in vivo* measurements understanding that an "appropriate blank" can be estimated from count results from an unexposed population of people.

Once the empirical decision level for a specific analytical process has been established, reporting requirements and processes need to be worked out. Section A.8 of ANSI N42.2 provides recommendations regarding results reporting. Our reporting is consistent with ANSI N42.2; however, we have specific recommendations regarding when and how this information should be reported. Our recommendation is to store the final ABACOS PLUS report in the radiation health record, including the following information recommended in ANSI N.42.2:

- sample identification code (i.e. name and social security number)
- reference date/time (specifically the count date/time)
- identification of the specific measurement procedure (counting system [LELM, HELM, WBC] and we would add key instrumentation information like make, model, serial numbers, etc.)
- identification of radionuclides specified for analysis and others found,
- the result reported as:
    - < Decision Level (Value & Units), if less than decision level, or
    - Result $\pm 2\sigma$ error, if greater than decision level.

We do not recommend the actual analytical result being stored in the radiation health record unless the result exceeds the empirical decision level. Storing data less than the empirical decision level only serves to confuse the record system over time. Restating what we said earlier in this paper, results below the decision level indicate that the subject is indistinguishable from the unexposed population from a bioassay standpoint, and followup is therefore not warranted.

Modifying the ABACOS PLUS Preliminary and Final Reports

It is necessary to modify the Canberra ABACOS PLUS software reporting process in order to implement use of the empirical decision levels in screening and reporting. This requires utilizing a spare field in the detail portion of the nuclide library. The field allows entry of a decision level value (as determined by the user). If the field is left blank the software performs the vendor designed report process (i.e., MDA reporting). The value entered in the field in the nuclide library replaces the calculated MDA for comparison during processing of spectra. The analysis algorithm compares the entered empirical decision level for the nuclide of interest to the calculated activity for the nuclide in the analyzed spectrum. If the analysis yields activity less than the empirical decision level value entered, then the decision level value entered in the field is reported, including the comment "Decision Level Reported" (see Figure 1). (The comment field of the final report is modified to contain the words "Decision Level Reported".) If a result greater than the empirical decision level value is calculated, then the calculated activity is reported (see Figure 2). This process coincides with result reporting recommendations in ANSI N42.2. However, for a result greater than the decision level, a recount for a longer period of time is performed to improve counting statistics, and if the recount is negative, the first result is commented in the record to this effect. If activity continues to be found, a review is started.

Because the user definable parameters operate on the full spectra, not just the region of interest, occasionally nuclides not of interest are also "found". However, no decision level has been determined for them because they are not of interest. These nuclides of no interest are easily

eliminated from the final report by reanalyzing the spectra using the vendor recommended values for MDA reporting. The nuclide of interest is still reported as less than decision level, if it in fact is, and nuclides not of interest are eliminated from the report, unless the calculated activity exceeds the calculated MDA, which is rare. Very rarely, a spectra has a nuclide of interest greater than the decision level and another nuclide not of interest that is identified as greater than MDA. On these rare occasions the results not of interest must be handled administratively. For example, occasionally when performing LELM uranium monitoring, uranium above decision level and Cs-137 are found. A recount on the LELM is performed, and the uranium is now determined to be less than decision level, however, the Cs-137 is still positive. The spectrum is reanalyzed to determine if the Cs-137 level exceeds MDA. If it does, the individual is counted on the Whole Body Counter, for which a Cs-137 decision level has been established. This is the result of forcing the ABACOS PLUS software to flag lower levels than normally programmed.

According to Canberra Industries, a new commercial version of ABACOS PLUS is scheduled for release in the Fall of '95 which will include decision level reporting capabilities (evidence that this type of reporting is becoming more and more, the method of choice). Hopefully this minor problem will be resolved in this software release.

DISCUSSION

The empirical decision level activity determined for Cs-137 whole body counts was 2.7 nCi. This should represent the 95th percentile of results, and when used as a screening tool, one that yields about a 5% false positive rate. A review of operational counting results since January 1, 1995 indicates 263 whole body counts were performed during this period. Fourteen of 263 (5.3%) were reported as "positive" on the initial count. All results were determined to be less than 2.7 nCi after a 15 minute recount. Results are consistent with the expectation that the calculated decision level should yield 5% false positives. The data indicates that by studying the standard deviation of the blank, as recommended by ANSI N13.30, and applying decision level theory, results meet the expectation. About 5 out of 100 results would be falsely determined to be positive.

This same process was used to determine a cobalt-60 empirical decision level using the ACCUSCAN II in the stationary mode for a 10 minute lung count. The empirical decision level was determined to be 0.5 nCi. Follow-up rates since January 1, 1995 are about 4.4%. All results were determined to be less than 0.5 nCi after a 20 minute recount.

Similarly, empirical decision levels for U-235 and Th-234 were determined for a 30 minute lung scan using the LELM. The empirical decision level for U-235 is 0.068 nCi, and for Th-234 is 0.62 nCi. Followup rates have been similarly reasonable, with all recounts determined to be less than the decision levels after a minimum of a 40 minute recount.

Continued Study of Non-radiological Worker and Radiological Worker Populations

A follow-up rate significantly above 5% would require an investigation into the cause, including the possibility of a radiological situation requiring remediation or anomalies in statistics obtained by counting the unexposed population. Follow-up rates much less than 5% may indicate that the empirical decision level requires adjustment downward. Either could be caused simply by a change in the analytical process that needs to be reviewed and corrected. The statistical test

offered at the beginning of the paper can be helpful in analyzing the data to assess the equivalency of the exposed and unexposed populations. In order to do this analysis, it is necessary to have the analytical result and its 2 σ error. We recommend a separate report which is provided at the same time as the results reports for radiation health records. This raw data is maintained in a file separate from the radiation health records, should its recovery be required.

Summary

Empirical decision levels provide a simple but powerful method of screening radiological worker *in vivo* sample results. The methodology is easily transferrable to *in vitro* and other analyses.