DOCKET NO: A-93-02
V-B-26

TECHNICAL SUPPORT DOCUMENT FOR

SECTION 194.55: COMPLIANCE ASSESSMENT STATISTICS

U. S. ENVIRONMENTAL PROTECTION AGENCY
Office of Radiation and Indoor Air
Center for the Waste Isolation Pilot Plant
401 M. Street, S. W.
Washington, DC 20460

MAY 1998

Technical Support for this document was provided by TechLaw, Inc. and its subcontractors under EPA Contract 68D50174.

TABLE OF CONTENTS

Tables

1. INTRODUCTION

Conservative estimates of groundwater pathway doses were projected by the U.S. Department of Energy (DOE) for a person residing next to the Waste Isolation Pilot Plant (WIPP) as part of the WIPP Compliance Certification Application (CCA) (DOE, 1996; DOE, 1997a). Additional dose projections were prepared for the Performance Assessment Verification Test (PAVT) mandated by the U.S. Environmental Protection Agency (EPA) (DOE, 1997b). Independent calculations were made in the Dose Verification Evaluation (DVE) to confirm the DOE dose modeling conducted for the CCA and PAVT (EPA, 1997). This report develops and applies statistical tests to verify that the doses and concentrations meet the requirements of §194.55(d), (e), and (f): In summary, §194.55(d) requires that sufficiently many estimates were generated, §194.55(e)requires that the full ranges of estimated doses and concentrations were displayed, and §194.55(f) requires that the mean and median estimated doses and concentrations do not exceed the maximum allowable dose or concentration, with a 95 percent confidence. Section 2 of this report discusses the methodology used in the verification procedure. Section 3 presents the results of the statistical analyses.

2. METHODOLOGY

Part §194.55 of Title 40 of the Code of Federal Regulations (CFR) details the requirements for the results of the individual and ground-water protection compliance assessments. In the CCA, a Latin Hypercube Sampling method was used to sample uncertain disposal system parameter values, in three replications of 100 realizations each. These same 300 realizations were used for the §191 and §194 compliance assessments. In the PAVT, a further set of 300 realizations, in three replications of 100 realizations, were generated.

To assess compliance with the individual and ground-water protection requirements, DOE estimated the following for each realization, assuming undisturbed performance:

(a) the maximum annual committed effective dose from all pathways for 10,000 years after disposal,

(b) the maximum total radioactivity level for radium 226 and radium 228 in any underground source of drinking water (USDW) for 10,000 years after disposal,

(c) the maximum gross alpha particle radioactivity level (including radium 226 but excluding radon and uranium) in any USDW for 10,000 years after disposal, and

(d) the maximum annual dose equivalent to the total body or any internal organ from beta particle and photon radioactivity in any USDW for 10,000 years after disposal.

In each of these four cases, the DOE estimated the dose or concentration using a bounding analysis based on unrealistic assumptions that would result in the overestimation of the dose or concentration. In the CCA, for 291 realizations out of 300, each of these four values were zero. In the PAVT, each of these four values were zero for 285 realizations out of 300. Sections 2.1-2.3 describe the methodology used to evaluate the criteria in §194.55(d), (e), and (f), respectively.

2.1 SUFFICIENCY OF THE NUMBER OF ESTIMATES

Under §194.55(d), the number of estimates generated shall be large enough such that the probability is at least 0.95 that the maximum estimate exceeds the 99th percentile of the population of estimates. The probability is 0.99 that a single estimate is less than the 99th percentile. Therefore, if the 300 realizations were statistically independent, the probability that the maximum estimate exceeds the 99th percentile of the population of estimates would equal $1 - (0.99)^{300} = 0.951$, and the §194.55(d) criterion would be satisfied. For the Latin Hypercube Sampling (LHS) method used in both the CCA and PAVT, the 300 individual estimates are not exactly independent, but are approximately independent. Furthermore, the LHS method is designed to be more likely to cover the range of possible parameter values than simple random sampling. On that basis, the probability that the maximum LHS estimate exceeds the 99th percentile of the population of estimates should also exceed 0.95, and the §194.55(d) criterion is satisfied.

2.2 COVERAGE OF THE RANGES OF DOSES AND CONCENTRATIONS

Under §194.55(e), the full ranges of estimated doses and concentrations (for cases a to d above) must be displayed.

Consider first the CCA. For case a, the all pathway individual dose, the USDW estimated doses were reported in Table 8-2 of the CCA and the maximum estimated dose (0.46 mrem/year) from other (indirect) pathways was reported in the DOE document DOE, 1997a. For case b, a bounding analysis in the CCA report shows that the maximum estimated radium concentration is 2.0 pCi/L across the nine non-zero realizations. For case c, Table 8-1 of the CCA contains the 300 estimated concentrations for the five radionuclides ^{241}Am, ^{239}Pu, ^{238}Pu, ^{234}U, and ^{230}Th. The nine non-zero ^{226}Ra concentrations were not separately reported, but the maximum gross alpha particle concentration was reported as 7.81 pCi/L. The confidence interval analysis described below used a less conservative approach that added the total radium concentration bound (2.0 pCi/L) to the total of the five radionuclide concentrations. For case d, the USDW dose, the 300 dose estimates were reported in Table 8-2 of the CCA. In the EPA's DVE, independent estimates were made and the all pathway and USDW doses were tabulated, either assuming no (radionuclide) daughter ingrowth or assuming 10,000 years of daughter ingrowth. Thus, the requirements of 194.55(e) have been satisfied for the CCA either by displaying specific estimates for each of the non-zero realizations or by displaying the same upper bound estimate for those realizations.

Similar reports were generated in the PAVT: For case a, the all pathway individual dose, the USDW estimated doses were reported in Table 3 of the PAVT and the maximum estimated dose (0.031 mrem/year) from other (indirect) pathways was reported in Table 5 of the document DOE, 1997c. For case b, the bounding analysis in Section 4 of the PAVT shows that the maximum estimated radium concentration is 0.49 pCi/L across the fifteen non-zero realizations. That analysis also corrected the CCA value from 2.0 pCi/L to 0.14 pCi/L (Note that since the statistical criteria in §194.55(f) are shown below to be satisfied for the CCA when the 2.0 pCi/L bound was used, these criteria must also be satisfied if the revised, lower radium concentration bound had been used.) For case c, Table 1 of the PAVT contains the 300 estimated concentrations for the five radionuclides ^{241}Am, ^{239}Pu, ^{238}Pu, ^{234}U, and ^{230}Th. The fifteen non-zero ^{226}Ra concentrations were not separately reported, but the maximum gross alpha particle concentration was reported as 2.4 pCi/L. The confidence interval analysis described below used a less conservative approach that added the total radium concentration bound (0.49 pCi/L) to the total of the five radionuclide concentrations. For case d, the USDW dose, the 300 dose estimates were reported in Table 3 of the PAVT. In the DVE, independent estimates were made and the all pathway and USDW doses were tabulated, either assuming no (radionuclide) daughter ingrowth or assuming 10,000 years of daughter ingrowth. Thus the requirements of 194.55(e) have been satisfied for the PAVT either by displaying specific estimates for each of the non-zero realizations or by displaying the same upper bound estimate for those realizations.

2.3 METHODOLOGY FOR VERIFYING THAT MEAN AND MEDIAN ESTIMATED DOSES AND CONCENTRATIONS DO NOT EXCEED MAXIMUM ALLOWABLE VALUES

Under §194.55(f), the compliance assessment shall document that the mean and median estimated doses and concentrations do not exceed the maximum allowable dose or concentration, with a 95 percent confidence. For cases a-d above, these maximum values are (a) 15 mrem/year, (b) 5 pCi/L, (c) 15 pCi/L, and (d) 4 mrem/year. The bounding analyses in the CCA and PAVT indirectly verified these requirements by showing that the maximum estimated dose or concentration was always lower than the maximum allowable value. The methodology for a direct verification using 95 percent confidence intervals for the means and medians is given in this section and the results are shown in Section 3.

Confidence Intervals for the Means

In each case, two alternative 95 percent confidence intervals for the mean were computed. The first approach assumes that the 300 realizations are approximately independent, and uses the central limit theorem of statistics to verify that the sample mean (of the 300 values including the 291 CCA zeroes or the 285 PAVT zeroes) is approximately normally distributed. The well-known 95 percent confidence interval for the mean is the sample mean plus or minus 1.96 times the sample standard deviation divided by the square root of the sample size, and the §194.55(f) requirement is met if the upper bound is no more than the maximum allowable value. The formulae are:

N	=	Sample Size = 300
X_i	=	Estimate for realization I (I= 1, 2, 3, ...300)
MEAN	=	ΣX_i / N
SD	=	$\{\Sigma(X_i - \text{MEAN})^2 / (N - 1)\}^{0.5}$
CONFIDENCE INTERVAL	=	MEAN $\pm$ 1.96 SD / $\sqrt{N}$

REQUIREMENT: MEAN + 1.96 SD/$\sqrt{N}$ <= Maximum Allowable Value.

Negative lower bounds are replaced by zero, since the mean must be nonnegative.

The second approach avoids the assumption that the LHS estimates and approximately independent, and instead assumes only that the central limit theorem can be applied to show that the three independent replication averages are approximately normally distributed. Since 4.303 is the 97.5th percentile of a t distribution with 3-1=2 degrees of freedom, this gives the alternative formulae:

K	=	Number of Replications = 3
Y_i	=	Sample mean for replication I (I= 1, 2, 3)
MEAN(Y)	=	ΣY_i / K
SD(Y)	=	$\{\Sigma(Y_i - \text{MEAN}(Y))^2 / (K - 1)\}^{0.5}$
CONFIDENCE INTERVAL	=	MEAN(Y) ± 4.303 SD(Y) / √K

REQUIREMENT: MEAN(Y) + 4.303 SD(Y)/√K <= Maximum Allowable Value.

Negative lower bounds are replaced by zero, since the mean must be nonnegative.

Confidence Intervals for the Medians

The 95 percent confidence intervals for the population median are derived as follows (see David, 1991). The dose or concentration is zero with some probability p, and is non-zero with probability 1-p. The nonzero values can be assumed to have a continuous distribution. The population median equals zero if and only if p >= 0.5.

The first approach assumes the 300 estimates are independent. When p < 0.5, the population median is positive, and the probability that a single estimate is less than or equal to the population median equals exactly 0.5. Therefore, the probability Prob(M) that the population median lies between the r'th and s'th highest estimated value (counting r=1 as the minimum) equals the probability of between r and s-1 "successes" in 300 independent binomial trials, where a success is a value less than or equal to the population median, and the success probability is 0.5. Symmetrically choosing s=301-r, Prob(M) is at least 0.95 if p < 0.5, r=133, and s=168.

When p >= 0.5, the population median equals zero, and the probability that a single estimate is less than or equal to the population median equals p. Therefore Prob(M) equals the probability of r or more "successes" in 300 independent binomial trials, where a success is a zero estimated value and the success probability equals p. When p >= 0.5, this probability is obviously at least as large as when the success probability equals 0.5. Also, the probability of at least r successes is obviously no more than the probability of between r and s-1 successes (if s > r). Hence, Prob(M) is at least 0.95 if p >= 0.5, r=133, and s=168.

Putting the results from the last two paragraphs together, it follows that the interval from the 133rd to the 168th highest value is a 95 percent or greater confidence interval for the population median based on the 300 realizations. The requirement is met if the 168th highest value is less than the maximum allowable value. (For the CCA and PAVT, this value was 0 in both cases, so the requirement was clearly met).

An alternative approach for the median analysis does not assume that the LHS estimates are statistically independent for the same replication. Let Y_1, Y_2, and Y_3 be the proportions of zero estimates for the three independent replicates. Assuming these Y_i values are approximately normally distributed, a 95 percent confidence interval for p is

$$\text{MEAN}(Y) \pm 4.303\ \text{SD}(Y) / \sqrt{3}.$$

If the lower bound of this confidence interval is greater than 0.5, then the $194.55(f) requirement will be met: Since we are 95 percent confident that $p >= 0.5$, we are at least 95 percent confident that the population median equals 0. In this case, all the median confidence intervals are from 0 to 0 and do not exceed the maximum allowable value.

3. VERIFICATION THAT MEAN AND MEDIAN ESTIMATED DOSES AND CONCENTRATIONS DO NOT EXCEED MAXIMUM ALLOWABLE VALUES

For each case, the lower and upper bounds of two alternatively derived 95 percent confidence intervals for the mean and median using the methodologies given in Section 2.3 were computed. Since the upper bound is lower than the maximum allowable value in every case, the §194.55(f) requirements are met.

3.1 MEANS

The results of the analyses of the means are summarized in Table 1 for the CCA and Table 2 for the PAVT. For each case we report the lower and upper bounds of the two alternative 95 percent confidence intervals and show that the upper bound is lower than the maximum allowable value, so that the §194.55(f) criterion is met.

3.2 MEDIANS

For both the CCA and the PAVT, assuming the 300 realizations are independent, a 95 percent or greater confidence interval for the median is from the 133rd to the 168th highest values, i.e. from 0 to 0. Since the maximum allowable value exceeds the upper bound of 0, the §194.55(f) criteria for the medians are met.

Assuming only that the proportion of non-zero estimates for each replication is approximately normally distributed, the alternative analysis also shows that the median equals zero with 95 percent or greater confidence, which again verifies that the §194.55(f) criteria for the medians are met:

Letting Y_1, Y_2, and Y_3 be the proportions of zero estimates for the three independent replicates and assuming these Y_i values are approximately normally distributed, a 95 percent confidence interval for p is

$$\text{MEAN}(Y) \pm 4.303\ \text{SD}(Y) / \sqrt{3}.$$

Table 1. CCA Verification. Confidence intervals for means A) assuming 300 independent realizations, and B) assuming normally distributed replication averages.

Description	Mean	95 percent confidence interval A	95 percent confidence interval B	Maximum Allowable Value
All pathways dose (mrem/yr): CCA Table 8.2 + 0.46 (DOE, 1997a)	0.017	0.005,0.028	0.000,0.036	15
All pathways dose (mrem/yr): No daughter ingrowth DVE Table 3-3	0.004	0.000,0.009	0.000,0.012	15
All pathways dose (mrem/yr): 10,000 years ingrowth DVE Table 3-3	0.004	0.000,0.009	0.000,0.013	15
Radium concentration (pCi/L): 9 values of 2.0 (CCA) 291 values of 0.0	0.060	0.021,0.099	0.000,0.159	5
Gross alpha concentration (pCi/L): CCA Table 8.1 + 2.0 (2.0 = maximum total radium concentration)	0.100	0.020,0.181	0.024,0.176	15
USDW dose (mrem/yr): CCA Table 8-2	0.003	0.000,0.007	0.000,0.009	4
USDW dose (mrem/yr): No daughter ingrowth DVE Table 3-1	0.003	0.000,0.007	0.000,0.009	4
USDW dose (mrem/yr): 10,000 years ingrowth DVE Table 3-1	0.003	0.000,0.007	0.000,0.010	4

Table 2. PAVT Verification. Confidence intervals for means A) assuming 300 independent realizations, and B) assuming normally distributed replication averages.

Description	Mean	95 percent confidence interval A	95 percent confidence interval B	Maximum Allowable Value
All pathways dose (mrem/yr): PAVT Table 3 + 0.031(DOE, 1997c)	0.0017	0.0008,0.0026	0.0010,0.0023	15
All pathways dose (mrem/yr): No daughter ingrowth DVE Table 3-4	0.0002	0.0000,0.0005	0.0000,0.0006	15
All pathways dose (mrem/yr): 10,000 years ingrowth DVE Table 3-4	0.0012	0.0000,0.0032	0.0000,0.0053	15
Radium concentration (pCi/L): 15 values of 0.49 (PAVT) 285 values of 0.0	0.025	0.012,0.037	0.012,0.037	5
Gross alpha concentration (pCi/L): PAVT Table + 0.49 (0.49 = maximum total radium concentration)	0.027	0.013,0.040	0.017,0.037	15
USDW dose (mrem/yr): PAVT Table 3	0.0002	0.0000,0.0004	0.0000,0.0005	4
USDW dose (mrem/yr): No daughter ingrowth DVE Table 3-2	0.0002	0.0000,0.0004	0.0000,0.0005	4
USDW dose (mrem/yr): 10,000 years ingrowth DVE Table 3-2	0.0009	0.0000,0.0024	0.0000,0.0039	4

For the CCA, $Y_1 = 0.99$, $Y_2 = 0.95$, and $Y_3 = 0.97$. The mean equals 0.97 and the confidence interval is (0.92, 1.02). Since the lower bound of this confidence interval is greater than 0.5, the $194.55(f) requirement is met: Since we are 95 percent confident that $p >= 0.5$, we are at least 95 percent confident that the population median equals 0.

For the PAVT, $Y_1 = 0.96$, $Y_2 = 0.94$, and $Y_3 = 0.95$. The mean equals 0.95 and the confidence interval is (0.93, 0.97). Since the lower bound of this confidence interval is greater than 0.5, the §194.55(f) requirement is also met for the PAVT.

4. REFERENCES

David, H.A., "Order Statistics", Wiley, 1981.

DOE, 1996, "Title 40 CFR Part 191 Compliance Certification Application for the Waste Isolation Pilot Plant," U.S. Department of Energy, DOE/CAO-1996-2184, October 1996.

DOE, 1997a, "Analysis Report for Estimating Dose from Cattle, Vegetable Consumption, and Inhalation Pathways Utilizing Contaminated Water from the Top of the Salado, Culebra, and Selected Marker Beds for an Undisturbed Case Supporting Review Compliance Certification Application," Document Version 1.01, Analysis Package in Support of the WIPP CCA. Dose Calculations of Other Pathways. WPO#43298.

DOE, 1997b, "Summary of the EPA-Mandated Performance Assessment Verification Test Results for the Individual and Groundwater Protection Requirements," PAVT Groundwater Protection. WPO#47258.

DOE,1997c, "Summary of the EPA-Mandated Performance Assessment Verification Test Results for Individual Protection Requirements: Estimated Doses to Internal Organs and Total Body from Groundwater Ingestion and to the Total Body from Beef Consumption, Vegetable Consumption, and Inhalation of Soil." WPO #47309.

EPA,1997, "Technical Support Document for Sections 194.51, 194.52, and 194.55: Dose Verification Evaluation," September 1997. Docket: A-93-02, V-B-25.