962

Filed: January 31, 1986.

UNITED STATES OF AMERICA

NUCLEAR REGULATORY COMMISSION

before the

ATOMIC SAFETY AND LICENSING BOARD

In the Matter of TEXAS UTILITIES GENERATING COMPANY et al. (Comanche Peak Steam Electric Station, Units 1 and 2)	Docket Nos. 50-445 OL 50-446 (Application for an Operating License)

APPLICANTS' MEMORANDUM IN RESPONSE
TO BOARD'S MEMORANDUM
(Statistical Inferences from CPRT Sampling)

Introduction

The Applicants submit this memorandum in response to the "MEMORANDUM (Statistical Inferences from CPRT Sampling)" issued by the Licensing Board on November 11, 1985 ("Memo"). In that memorandum the Board noted several questions and "concerns" that were apparently raised during the Board's review of the Program Plan

D503

published by the Comanche Peak Response Team ("CPRT") Senior Review Team ("SRT") on June 28, 1985 (Revision 2), as well as by the Board's review of the transcripts of several meetings at which aspects of the Program Plan were discussed.

Neither the CPRT Program Plan nor any of the comments that the Board may have read are in evidence in these proceedings. Indeed, the Program Plan and several of its constituent attachments, appendices and action plans have been modified several times since the publication of Revision 2, and further revisions are possible (if not likely). Given that nothing is yet before the Board, we assume that the statements contained in the Board's memorandum are neither findings nor conclusions. See Memo at 3 n.3. The purpose of this memorandum is not to serve as testimony or evidence, but rather to indicate to the Board that the Applicants and the SRT have the Board's questions in mind.

This response is structured in parallel with the comments in the Board's memorandum, as follows:

1. The Role, Use and Design of Sampling in the CPRT Program;

2. The Relationship Between Problems and their Resolution;

3. Stratification of Samples;

4. The "Level of Safety Chosen" by the CPRT and Relationship to Appendix B and FSAR Commitments; and

5. Exemption from Appendix B.

1. The Role, Use and Design of Sampling Studies in the CPRT Program

The role of sampling in the CPRT Program is to serve as one of several approaches for adding rigor and consistency to final judgments of the CPRT. Sampling can also help CPRT assure that a systematic process is used in the search for potential deficiencies and the detection of potential programmatic weaknesses and potential root causes. The CPRT Program does not use sampling alone to draw aggregate conclusions about design, construction or testing activities.

Consistent with the CPRT Sampling Policy, which is set forth in Section 1.0 of Appendix D to the Program Plan (Rev. 1, 1/30/86) (a copy of which is attached for the convenience of the Board), action plans are not formulated to rely solely on sampling studies or to result in purely statistical statements; action plans, rather, use sampling techniques to augment analysis, knowledge of plant conditions and testing results, knowledge of program and procedure provisions, and knowledge of inspector or other personnel qualifications and performance and engineering judgment. The following information is

included, either directly or by reference to Appendix D of the CPRT Program Plan, in each action plan that involves the use of random sampling:

1. Population definition, including any stratification (to achieve homogeneity), or any engineering bias relative to the sample selection.

2. A statement of the statistical "screen" to be applied (i.e., the minimum deficiency rate that will be detected at the stated confidence level).

3. A statement of the accept/reject criteria, including the relevant test statistics and critical region.

Computations will be summarized, and conclusions will be presented in the results reports.[1]

---

[1]The Board quotes from Introduction to Statistical Analysis (4th ed., 1983) at 85, where the authors make some general statements regarding statistical studies in the context of a specific discussion of the hypothesis on the mean test. The SRT has no quarrel with those general statements; it points out, however, that more commonly employed in the ISAPs employing statistical analysis is a test for the existence of a non-parametric population extreme, in this case the existence of a safety-related construction or design deficiency. This is a type of analysis described, briefly, in the same reference at 401:

> "In quality inspection it is sometimes desired to guard against or to detect individuals at only one extreme of the observational scale. Table A-25e gives samples sizes corresponding to

2. The Relationship Between Identified Problems and their Resolution

Once a potential problem area has been identified by either an external or internal source, a plan of action is formulated to confirm or rebut the existence of a programmatic problem and to resolve any actual problems encountered. Tasks are identified in the action plan that, when completed, will provide reasonable assurance that a programmatic problem does not exist, or if one is detected, that it is properly dispositioned. If any of these tasks require

---

> the probability . . . of finding [a given] proportion . . . of the population at the specified end of the distribution; thus the single extreme provides one-sided tolerance limits."

Id. Non-parametric tests are often used in engineering evaluations of construction quality because the reliability of the statistical inference does not depend upon the validity of any a priori assumption about population distribution:

> "The level of significance for a non-parametric method is not affected by the population distribution. Therefore the experimenter can be sure of knowing the chance of rejecting the null hypothesis when it is true."

Id. at 408.

reinspections and/or documentation reviews within a particular population, the CPRT may deem sampling and inference to be an appropriate investigation tool.

Before any sampling is done, the objective of the sampling must be clearly stated with a logical relationship to the action plan objective. The appropriate population or stratum of items or quantity of material is identified. This population must be reasonably homogeneous with respect to the parameter or attribute thought to be the problem. Data are then gathered from the population of items, the nature of which (e.g., a particular population parameter) is unknown. These data are used directly to infer the value of the unknown parameter. In most cases the parameter is the percentage of the population which is deficient and the inference will be statistical. For most of the ISAPs, the CPRT has adopted a 95/95 one-sided upper-bound confidence limit or probability interval screen on the percentage deficient in order to make judgments about the existence of programmatic

problems.[2] If the 95 percent upper-bound confidence limit (or probability interval) calculated from the sample is equal to or greater than 5 percent, the population of items is said to have failed the screen and further investigation is necessary.

The statistical inference approach used in the CPRT effort can be related to hypothesis testing, though there is a large (perhaps infinite) number of hypotheses that could be associated with the upper-bound confidence limit approach. One hypothesis test is based on the assumption that any of the programmatic

---

[2]The "95/95" test (which is sometimes referred to as a "95/5" test, the two being interchangeable in this context) is believed to be one commonly employed in nuclear cases and has been accepted as providing the requisite reasonable assurance. See, e.g., Public Service Electric and Gas Co. (Salem Nuclear Generating Station, Unit 1), ALAB-650, 14 NRC 43, 54-55 (1981):

> "Despite the Colemans' skepticism, Exxon's guaranty of 95 percent leak-tightness with a 95 percent confidence level amply satisfies the public health and safety standard of the Atomic Energy Act."

(Footnote omitted.)

type of problems asserted to have existed with the CPSES QA/QC program would be expected to result in a deficiency rate of at least 5% or more, if it existed in fact; one would then have the null hypothesis state that the population deficiency percentage is less than 5 percent. The alternative hypothesis is that the percentage of deficient items in the population is equal to or greater than 5 percent. The level of significance would be 0.95. The test statistic for this particular hypothesis test is the number of deficient items found in a sample of 60, and the critical region is one or more deficient items in the sample.

Where this form of statistical investigation is employed, however, the CPRT is not attempting to test any particular hypothesis. Rather, it is using a confidence limit (or probability interval) statistical inference approach to screen potential programmatic problems in any given population. If a population passes the 95/95 sample screen, the CPRT is quite confident that the population of items is free from programmatic deficiencies. (Passage also indicates a low level of non-programmatic (or random) errors, as

well.) If a population does not pass the 95/95 sample screen, the deficiency rate is considered high enough for the CPRT to be concerned that a programmatic problem may exist, and that further evidence is needed to verify that an actual problem exists and, if it does, to determine the extent of the problem. This signals the CPRT that further investigation is needed to determine the root cause and generic implications of such a deficiency. This would lead to sample expansion as discussed in Attachment 4 to Appendix D.

3. Stratification of Samples

The CPRT is indeed structuring its investigations carefully with regard to deficiencies and adverse trends of deviations detected or suggested by either an external or internal source. These identified problems, which are considered to be potential programmatic deficiencies, are pursued with more rigor than an initial self-initiated exploratory sampling. Generally, after evaluating a deficiency or a trend of deviations, a root cause will be hypothesized to exist in the entire population or a subset (stratum) of the population.

If a deficiency or adverse trend of deviations is detected or suggested by an external source or in the CPRT sample, and no root cause associated with a particular stratum of the population can be isolated, sample expansion into the entire population will continue until it is determined that either the deficiency was a random occurrence of very low frequency (i.e., no programmatic problem exists), or a potentially programmatic deficiency is identified in a stratum. If deficiencies continue to be detected, or two or more deficiencies are detected in the initial minimum CPRT sample and they cannot be associated with a specific root cause (stratum), 100 percent of the population will be inspected or reviewed with all deficiencies detected and dispositioned.

The CPRT requires sampling (either initial or expanded) in a stratum when it has been hypothesized that the stratum contains and bounds the deviations or deficiencies of the type already detected. Such a stratum may be suggested by external sources, by a deviation trend in the initial CPRT sample, or by a root cause evaluation originating inside or outside the population of items being inspected or reviewed. In

addition, unless otherwise justified, sample expansion in the general population, minus the identified stratum, is required to verify that the deficiency is not associated with the remaining strata.

In general, the CPRT has recognized the importance of stratification on the overall investigation effort and has spent considerable time and effort in identifying and defining homogeneous strata in order to enhance the inference process (see Appendix D for further discussion of stratification). In a nuclear power plant, as with all complex systems, stratification into homogeneous populations for sampling investigations must be based on the particular information sought. Note that if carried to the extreme, this process could inadvertently end up with single items being contained in a given stratum. Therefore, from a practical engineering point of view, certain characteristics are considered unimportant in determining the boundaries of a homogeneous stratum. If, however, after a sample is found to contain deficiencies or an adverse trend of deviations asscciated with items possessing a certain characteristic and not with items that do not possess

this characteristic, further stratification would be carefully considered.

As this discussion reveals, unstratified samples are not intended to be persuasive regarding the non-existence of already found (or subsequently found) deficiencies. In cases where programmatic deficiencies have been or may be found, corrective action will be taken, often involving 100% reinspection of the potentially affected item. In cases where identified deficiencies are not programmatic, a root cause analaysis is generally combined with a focused statistical investigation to determine the nature and scope of the problem.

4. "The Level of Safety Chosen" by the CPRT and Relation to Appendix B and FSAR Commitments

The Senior Review Team respectfully does not concur in the Board's suggestion that the CPRT program poses any "hard question . . . about the level of safety that must be assured," if by this statement the Board is suggesting the need to "calibrate," with some sort of mathematical precision, the CPRT sample reinspection efforts to some assumed "level of safety" that would be expected to result from the implementation of an unquestioned QA/QC program per 10 C.F.R. App. B. The CPRT Program Plan is intended to stand on its own as a demonstration of a basis for reasonable assurance that, as designed and constructed, Comanche Peak Steam Electric Station is free from undetected and uncorrected safety signficant deficiencies that would preclude the facility from being operated safely.[3]

---

[3]If, in fact, the results reveal no safety significant deficiencies in the original or first-expanded sample in any particular area of

It is undeniable that absolute perfection of construction (or design) is neither achievable in connection with a project of this magnitude nor required.[4] A fortiori there is neither requirement nor assumption that a QA/QC program implemented in strict accordance with Appendix B will catch each and every construction (or design) error that might be made over

---

construction or design, then the implementation of the Program Plan will also provide assurance that the Appendix B QA/QC program in that area was free of the sort of inadequacy that would tend to permit the occurence of undetected safety significant construction or design deficiencies.

[4]The establishment and effective implementation of a QA/QC Program that is fully compliant with Appendix B can be expected to substantially contribute to the reduction of undetected and uncorrected defects. Due to the number and complexity of the activities associated with the design and construction of a nuclear power plant, however, it is impractical to expect that a "zero defect" standard (while a desirable goal) can be met. Accordingly, nuclear power plant design and construction is predicated on a philosophy of "defense in depth," which utilizes principles of diversity, redundancy and conservative safety margins. This philosophy assumes that defects can occur and, therefore, provides the capability to safely withstand the effects of such defects.

the course of the project[5] -- that is to say, discovery of a safety-significant design or construction deficiency does not, by itself, negate the existence or implementation of a fully compliant Appendix B program. Tne CPRT can and does make no claim of absolute perfection for itself, either. The Senior Review Team knows, however, of no way to quantify the level of detection of an Appendix B Program for random errors, to quantify the level of detection of either type of program for systematic errors (except to note that, qualitatively, the more wide-sweeping the hypothesized systematic error, the more closely the level of detection approaches 100%), or to "calibrate" one to the other.

---

[5] "We note that, in theory, a design quality assurance program will provide 100 percent review of the design work. The record is clear, however, that such a program can never assure that there will be no design errors. . . . Indeed, Appendix B only provides that the purpose of a quality assurance program is 'to provide adequate confidence that a [safety-related] structure, system or component will perform satisfactorily in service.'" Pacific Gas and Electric Co. (Diablo Canyon Nuclear Power Plant, Units 1 and 2), ALAB-763, 19 NRC 571, 593 n.86 (1984). There is, for this point, no valid distinction between design and construction.

CPRT is not simply a program of statistically-based sample reinspections. CPRT was formulated to respond to assertions of systematic or programmatic QA/QC breakdowns during the CPSES construction and design and to test, with reasonable assurance, the validity of the hypothesis that such a breakdown occurred. CPRT takes into account all that is known about program and procedure quality, inspector qualifications, the results of NDE and functionality testing and any other data that may be available about the extent to which the as-constructed hardware is free of safety-significant construction errors or design errors that will preclude achievement of the design bases set forth in the FSAR and applicable regulations and codes.

Exemption from Appendix B

The Applicants respectfully disagree with the Board's legal assertion (Memo. at 4) to the effect that, in the absence of a demonstration that sampling programs yield the same "level of safety" that is assertedly implicit in Appendix B, then "the plant would appear to fail to meet Appendix B requirements and to require the granting of an exemption from Appendix B pursuant to 10 CFR 2.758." Compliance with the construction QA/QC program required by Appendix B is not, either in terms or as applied, a condition precedent to the granting of an operating license. The only aspect of Appendix B that is a prerequisite to issuance of an operating license is the inclusion in the FSAR of "information pertaining to the managerial and administrative controls to be used to assure safe operation." 10 C.F.R. Part 50, Appendix B, "Introduction." Operational QA is not a subset of the sole remaining contention in these proceedings.

It is of course true that the failure to implement aspects of Appendix B applicable to the design and construction of the facility might be the grounds for enforcement action by the Staff, but no

"exemption" is required (even assuming such failures) in order for an operating license to issue. The only relevant licensing standard applicable to a construction QA/QC contention (such as Contention 5 in this case) is that set forth in 10 C.F.R. § 50.57(a)(3)(i) (i.e., that there is "reasonable assurance" or "adequate confidence" that the facility has been designed and constructed such that it is capable of being operated safely). See Pacific Gas and Electric Co. (Diablo Canyon Nuclear Power Plant, Units 1 and 2), ALAB-763, 19 NRC 571, 587-88 (1984), modified on a different point, CLI-84-84, 20 NRC 285 (1984).[6]

---

[6]"This argument [that no statistically valid error rates could be determined for the design sample reinspection there involved] essentially overlooks the standard by which the applicant's program is to be judged. We must determine whether the nonseismic verification program provides 'adequate confidence' that the nonseismic design of the safety-related structures, systems and components is proper so that such structures, systems and components will perform satisfactorily in service. This qualitative standard is not numerically quantifiable into expressions of probability or errors or error rates . . . .

"Pointing to the Commission's regulations, 10 C.F.R. § 50.57(a)(1), the Governor and the

See also Philadelphia Electric Co. (Limerick Generating Station, Units 1 and 2), ALAB-819, 22 NRC 681, 729 (1985); Louisiana Power & Light Co. (Waterford Steam Electric Station, Unit 3), ALAB-812, 22 NRC 5, 14-15,

---

joint intervenors repeatedly assert in their proposed findings that the applicant's verification program, in order to be sufficient, must demonstrate that the design of Diablo Canyon meets its license application requirements or licensing criteria. The application requirements and licensing criteria for Diablo Canyon, like any nuclear power plant, are spelled out in the various documents comprising of the operating license application including, most prominently, the applicant's Final Safety Analysis Reports (FSAR). The FSAR is a multivolume description of the entire facility containing literally thousands of so-called 'licensing criteria' ranging from safety significant ones to insignificant and extreme minor specifications or descriptions of details that have no safety implications. . . . In their proposed findings of fact, the Governor and the joint intervenors do not distinguish between safety significant and nonsafety significant licensing criteria. For example, the Governor and joint intervenors argue, relying on the staff's and the applicant's witnesses, that the nonseismic design does not meet licensing criteria because it is a virtual certainty that there remain undetected design errors in the unreviewed portions of the design. . . . But the witnesses relied upon by the Governor and joint intervenors all testified that not only was it likely there remained some design errors, but that it was extremely unlikely any of the errors were safety significant. . . . The central issue with respect to the proper design of Diablo Canyon, or any

44 (1985)[7]; Duke Power Co. (Catawba Nuclear Station, Units 1 and 2), ALAB-813, 22 NRC 59 (1985); Union

---

other facility, is the conformance of the design to the significant and substantive safety requirements and licensing criteria. To conclude otherwise would ignore reality and substitute 'perfection' for the regulatory standards of 'adequate confidence' and 'reasonable assurance.'" Id. at 587-88 & n.68.

[7]"LP&L, however, cannot turn back the clock and enlarge the QA staff that oversaw construction at the Waterford site. Our focus, then, must be on whether any significant construction deficiencies resulted and remain as a consequence of LP&L's inadequate staffing, and whether LP&L has taken steps to prevent understaffing in the future." Id. at 21. See also id. at 26 (trending inadequacies), 27 (audit inadequacies) ("There is little doubt that the implementation of LP&L's audit program was lacking and led to the 1982 QA breakdown. . . . Given these serious deficiencies, two pertinent questions arise. Are the auditing failures responsible for actual hardware or workmanship deficiencies that may remain unidentified and uncorrected? What assurance is there that these auditing failures will not recur?"), 31 (NCR deficiencies). Despite the failure to comply with various aspects of Appendix B, there was no suggestion by the Appeal Board that an exemption was required as a condition precedent to authorization of an operating license.

Instructive is the Appeal Board's resolution of the audit problem: "Audits are an important element of an applicant's overall QA program and are required by

Electric Co. (Callaway Plant, Unit 1), ALAB-740, 18 NRC 343, 346 (1983). All that is required is reasonable assurance that the as-constructed facility is constructed such that it is capable of being operated safely, and the results of the implementation of a construction Appendix B QA/QC program is but one way of providing evidence in support of that assurance. Sample reinspection programs are another accepted way of provided the same assurance. E.g., Waterford, ALAB-812, supra; Commonwealth Edison Co. (Byron Nuclear Power Station, Units 1 and 2), ALAB-793, 20 NRC 1591 (1984). In no case of a found or hypothesized breakdown of the construction QA program, however, has

---

10 C.F.R. Part 50, Appendix B, Criterion XVIII. Nonetheless, as the staff has explained, through the systematic sampling of various work and the QA documentation for it, audits provide but a third level of assurance. The principal levels of assurance are provided by, first, quality craftmanship and, second, quality inspections. . . . But as to [the work for which the audit program was determined to have been deficient], the first and second levels of assurance were, in fact, provided. This is demonstrated by the absence of significant safety deficiencies in the improperly audited Mercury work, as revealed by the major reinspection of that work that was undertaken by qualified personnel." 22 NRC at 28.

an exemption been required as a condition precedent to authorization of an operating license.

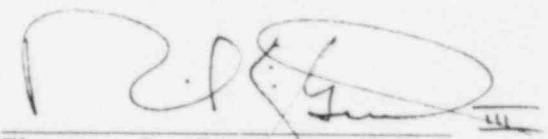
Respectfully submitted,

Robert A. Wooldridge
Worsham, Forsythe, Sampels & Wooldridge
2001 Bryan Tower - Suite 3200
Dallas, Texas 75201
Telephone: (214) 979-3000

Nicholas S. Reynolds
William A. Horin
Bishop, Liberman, Cook, Purcell & Reynolds
1200 Seventeenth Street, N.W.
Suite 700
Washington, D.C. 20036
Telephone: (202) 857-9800

Roy P. Lessy, Jr.
Morgan, Lewis & Bockius
1800 M Street, N.W.
Washington, D.C. 20036
Telephone: (202)872-5000

Thomas G. Dignan, Jr.
R. K. Gad III
Ropes & Gray
225 Franklin Street
Boston, Massachusetts 02110
Telephone: (617) 423-6100

By 
Thomas G. Dignan, Jr.
R. K. Gad III

Dated: January 31, 1986.

APPENDIX D

CPRT SAMPLING POLICY, APPLICATIONS AND GUIDELINES

1.0 POLICY STATEMENT

The Senior Review Team (SRT) has determined that, in general, it is unnecessary to examine an entire population of items or quantity of material in order to determine whether programmatic problems exist. By sampling a portion, inferences can be made regarding the entire population. The basis for CPRT decisions on design, construction, testing and QA/QC adequacy will be supported by sound engineering evaluation techniques, which often may include principles of sampling. Sampling and resulting inferences can be used as a powerful tool in identifying programmatic safety-significant deficiencies in programs and processes. Although the process of drawing inferences from sampling is not the sole means of reaching reasonable assurance that the plant design and construction are adequate, sampling may be a significant contributor to that evaluation.

It is also recognized that, for sampling to result in meaningful information about a process or program, the items in the population to be sampled must be similar (i.e., homogeneous) in the significant traits or attributes associated with that process or program. Some populations will be homogeneous by virtue of the work process by which they were made (e.g., ASME pipe support welding), others will be similar by virtue of the design activity that created them (e.g., containment isolation valve closure time), and so forth. Since sampling is utilized for a variety of purposes in the CPRT program it is essential that, when sampling is used, the population to be sampled is homogeneous and the objectives of the sampling are clearly tied to those of the action plan under consideration.

There are two basic ways to sample: one is to use judgment to select from a population those items that are likely to be the most critical, the other is to randomly select items from the general population (e.g., random selection of welds in a support structure). In the first method, called "biased sampling," the validity of one's inferences depend to a considerable extent on the validity of the investigator's prejudgment. In the second method, the samples are drawn randomly; and the resulting inferences depend on little or no bias from human prejudgment. Within the scope of the CPRT, both of these approaches are used to investigate various areas of interest and are justified within the context of their applications. In many cases, both approaches are used in the investigation of a single area of interest

APPENDIX D
(Cont'd)

2.0 APPLICATIONS

The purpose of this appendix is to:

- Delineate the various applications of sampling within the CPRT program.
- Set forth consistent guidelines for the mechanics of selecting samples wherever random sampling techniques are used in ISAPs and DSAPs (including the TRT issues, the Design Adequacy Program and Quality of Construction Program).

2.1 Quality of Construction (QOC)

The construction process produces hardware by execution of a number of relatively uniform construction activities. Therefore, the construction process is inherently susceptible to isolated hardware discrepancies. The overall frequency of deficiencies relative to the total number of opportunities is typically low, unless a programmatic problem exists.

To obtain a consistent sampling approach in the QOC Reinspection/Document Review (Issue-Specific Action Plan VII.c), described as the self-initiated investigation in Appendix B, the SRT believed that an initial sample screen should be based on a specific standard. The SRT has concluded that a 95/5 sample plan, when used in the context of homogeneous populations of attributes, would provide a reasonable screen to detect programmatic or systematic deficiencies*. Such a screen would ensure a sufficient initial sample size to evaluate the adequacy of the safety-significant attributes associated with each of the homogeneous work activities (HWAs) in the VII.c investigation. Accordingly, an initial random sample of at least 60 items is required for each homogeneous population (see Attachment 1, Table 1).

2.2 Other ISAPs

Many of the other ISAPs (i.e., TRT issues) utilize sampling techniques to investigate specific areas of concern. In general, the SRT requires that the sample sizes in each of these cases be consistent, at a minimum, with that required by the use of a minimum 95/5 sample screen. Any exceptions to this general principal are approved by the SRT, based on a case-specific review, and are reflected in the associated ISAP.

---

* A deficiency rate as low as 5% in a population will be detected by a 95/5 sampling plan with a probability or confidence level of 0.95.

APPENDIX D
(Cont'd)

2.0 APPLICATIONS (Cont'd)

2.3 Design Adequacy Program (DAP)

The focus of the Design Adequacy Program (DAP) is on the verification of the end products of the engineering and design process (i.e., designs represented by drawings, evaluations, or design specifications). In contrast to the construction process, where relatively few HWAs apply to large numbers of individual hardware items, the engineering and design process is characterized by a large number of homogeneous design activities (HDAs) with comparatively few design outputs being covered by each one.

The important aspect of the HDAs is that they include items for which a high degree of correlation exists in the design criteria, methodology, and procedures. Accordingly, evaluation of the adequacy of each HDA can be based on evaluating a representative selection of items within each HDA. The number of selected items will be sufficient to justify inferences and extrapolations that are appropriate for all items within each HDA. Attachment 4 of Appendix A to the CFRT Program Plan presents further details on the establishment of HDAs and the criteria for selecting items for evaluation.

If, in the event the DAP uses statistically-based sampling in the verification of any HDA, the sampling will be conducted in accordance with the provisions of this appendix.

3.0 GUIDELINES FOR RANDOM SAMPLING

The purpose of the attached guidelines is to:

- Assist in the development of non-parametric sample screens for Issue-Specific Action Plans (ISAPs) or Discipline Specific Action Plans (DSAPs) where random sampling is used (Attachment 1),
- Outline the use of one-sided tolerance limits for evaluating special cases of parametric attributes (Attachment 2),
- Outline the basic methods to be used in generating random samples from a population of items or attributes (Attachment 3),
- Outline the methods to be used for expanding samples (Attachment 4).

APPENDIX D
(Cont'd)

3.0 GUIDELINES FOR RANDOM SAMPLING (Cont'd)

These guidelines apply to sampling screens for most ISAPs and DSAPs. If other types of sampling applications arise, they must be considered on a case-by-case basis.

APPENDIX D
(Cont'd)

ATTACHMENT 1

GUIDELINES FOR SAMPLE INSPECTION OR REVIEW OF ATTRIBUTES

Table 1 of this attachment is generally used by CPRT to determine the sample sizes and corresponding detection numbers which are consistent with a 95 percent confidence level (or 0.95 probability) on the 5, 2.5, and 1 percent upper bound population percentage screens. Unless otherwise justified and specifically approved by the SRT, the number of deficiencies allowed in the sample screen will be no more than one (see Attachment 4 for discussion of sample expansion where one deficiency is identified). These sampling plans are based on the assumption of an infinite population size and are conservative when compared to sampling plans based on finite populations.

The minimum sample size for a 95/5 screen is 60 with a detection number of zero (i.e., the critical region is one or more detected items). This means that out of a random sample of 60 items inspected, if no items are found to belong to the classification of interest (e.g., deficient), there is a 95 percent confidence (or 0.95 probability) that less than 5 percent of the population will be in this classification. If items belonging to the classification of interest are detected in a minimum sample (i.e., the number detected is in the critical region), the 95 percent upper-bound confidence limit (or 0.95 probability interval) will be greater than 5 percent. It is still possible that the population percentage is less than 5 percent, but based only on the initial sample evidence, the probability that this is so is less than 0.95.

A root cause evaluation of the deficiency is performed in order to isolate a potentially deficient stratum from the population. If such a stratum is identified, sample expansion into that stratum is used to verify that indeed the deficiency is associated with the identified stratum. Sample augmentation in the remaining population (minus the potentially deficient stratum) is used to verify that the deficiency is not associated with the remaining stratum. Sample expansion is further discussed in Attachment 4 to this appendix.

APPENDIX D
(Cont'd)

ATTACHMENT 1
(Cont'd)

TABLE 1

SAMPLING PLANS FOR DETECTING UPPER-BOUND POPULATION PERCENTAGES (p_u) AT 95 PERCENT CONFIDENCE LEVEL*

SAMPLE SIZE**			DETECTION	CRITICAL
p_u = 5.0%	p_u = 2.5%	p_u = 1.0%	NUMBER	REGION
60***	120	300	0	1 or more
95	190	474	1	2 or more
126	252	630	2	3 or more
155	310	775	3	4 or more
183	366	915	4	5 or more
210	421	1051	5	6 or more

* Or 0.95 probability level.

** Sample sizes are determined from A. H. Bowker, and G. J. Lieberman, Engineering Statistics, 2nd Edition, Prentice-Hall, 1972, page 538. Note that these same sample plans may also be derived from Bayes' theorem, and are therefore applicable for sample expansion, using Bayes' theorem (see A. Boissonnade "CPRT Sampling Plans-Addendum," Civil/Structural/Mechanical CPRT File No. 11.1-005, or Box and Tiao, Bayesian Inference in Statistical Analysis, Addison-Wesley, 1973).

*** For populations of 100 or fewer items, the minimum sample size may be reduced to 45, with a detection number of zero. This is based on the hypergeometric distribution. Reference: Lieberman, and Owen, Tables of the Hypergeometric Probability Distribution, Standard University Press, 1961.

APPENDIX D
(Cont'd)

ATTACHMENT 2

SAMPLING GUIDELINES FOR ONE-SIDED TOLERANCE LIMITS

In some special cases ISAPs or DSAPs (or an evaluation of an adverse trend) may require the determination of a parametric tolerance limit of a particular attribute associated with items of a population. The acceptable quality of a population of items or quantity of material is often specified by setting a lower (upper) bound value based on the criterion that a certain percentage of the population fall above (below) this value (e.g., the concrete code specifies that at least 90 percent of the 28-day cylinder strengths fall above the required design strength). A lower (upper) bound population percentage is then inferred from a sample, compared with criterion value and the population either accepted as is, or corrective action taken. When a lower (upper) bound population percentage is specified in statistical terms, it is called a tolerance limit. A one-sided tolerance limit has the property that a certain percentage of the population of values (e.g., 90 percent) may be expected to fall above or below this bound with some level of confidence (e.g., 95 percent confidence).

A one-sided tolerance limit is defined as $\bar{X} - KS$ ($\bar{X} + KS$), where $\bar{X}$ is the sample average, and S is the sample standard deviation. The tolerance factor, K, is dependent upon the sample size, the specified population percentage above (below) the limit, and the desired level of confidence (e.g., the 95 percent confidence level). Once the confidence level has been selected and the population percentage specified, the sample size is only a function of the tolerance factor K. To lower the tolerance factor, it is necessary to increase the sample size. The relationship for several population percentages is listed in Table 2.

For ISAPs or DSAPs requiring the use of one-sided tolerance limits, sampling plans are developed by first determining, through engineering, materials, or other types of evaluations, that the underlying population distribution is either normal or log-normal*. Then, as a minimum, a sample size of 50 is obtained. The actual sample size selected, however, takes into account the difficulty in obtaining the sample and how sensitive the resulting conclusions are to the actual tolerance limit.

There is no unique sample size to be used for any particular tolerance limit problem. However, it is obvious from Table 2 that it becomes increasingly difficult to lower the tolerance factor as the sample size increases. From a practical point of view, sample sizes between 50 and 100 provide reasonable tolerance factors for the sampling effort.

---

* A goodness-of-fit test should be used to aid in evaluating the reasonableness of the assumed underlying distributions. Any tolerance limit applications for which the underlying population distribution cannot be reasonably assumed to be normal or log-normal will be handled on a case-by-case basis.

APPENDIX D
(Cont'd)

ATTACHMENT 2
(Cont'd)

TABLE 2

ONE-SIDED TOLERANCE LIMIT
FACTORS, K, FOR 95 PERCENT CONFIDENCE LEVEL

Sample Size	First* (ninety-ninth) Percentile	Fifth* (ninety-fifth) Percentile	Tenth* (nintieth) Percentile	Fiftieth** Percentile
5	5.75	4.21	3.41	0.90
10	3.98	2.91	2.36	0.56
15	3.52	2.91	2.36	0.45
20	3.30	2.40	1.93	0.38
25	3.16	2.29	1.84	0.34
30	3.06	2.22	1.78	0.31
35	2.99	2.17	1.73	0.28
40	2.94	2.13	1.70	0.26
50	2.86	2.07	1.65	0.24
70	2.77	1.99	1.58	0.20
100	2.68	1.93	1.53	0.17
300	2.52	1.80	1.42	0.10

* Reference: D. B. Owen, Handbook of Statistical Tables, Addison Wesley, 1962, page 126. Note that the first percentile means that 99 percent of the population falls above, one percent falls below.

** Reference: F. A. Webster, "Developing Sampling Plans for TRT Issues", Civil/Structural/Mechanical CPRT, File No. 11.1-001, 3/12/85.

APPENDIX D
(Cont'd)

ATTACHMENT 3

GUIDELINES FOR GENERATING RANDOM SAMPLES

The procedure for generating a random sample begins by first defining the unit to be sampled (e.g., truckloads of concrete, conduit runs, conductor terminations, etc.), then determining the total number of these units or items in the population. Note that the population, so defined, may actually be a subpopulation which has certain specified engineering attributes (i.e., a stratum). Each unit in the population (or stratum) must be assigned a unique sequential number 1 through N, where N is the total number of units. A table of random digits or a random number generator is then used to develop a random sequence of units from the population. Table 3 outlines the complete procedure.

APPENDIX D
(Cont'd)

ATTACHMENT 3
(Cont'd)

TABLE 3

PROCEDURE FOR GENERATING A RANDOM SAMPLE FROM A POPULATION

1. Determine population size, N, and number each item sequentially, 1, 2, ...N.

2. Start at a random position in a table of random digits or use a random seed in a random number generator and perform the following steps for each random five digit decimal fraction in sequence, until desired sample size is obtained.

3. If using a table of random digits, place a decimal point in front of each set of five digits* and multiply by the population size. If using a random number generator which produces five digit decimal fractions*, simply multiply by the population size.

4. Retain only the integer part of the above product and add 1. This will define the i^{th} item to be included in the random sample.

5. It is usually a good idea to generate a longer list of randomly selected items in case a particular item is inaccessible in the field, or in case the same item is selected more than once.

Example: Generate a sample of 300 items from a population of size 3791.

RN1	=	.04146**	x	3791	=	157.17486	=	157	+ 1	= 158
RN2	=	.23432	x	3791	=	888.30712	=	888	+ 1	= 889
RN3	=	.74381	x	3791	=	2,819.78371	=	2,819	+ 1	= 2,820
.		.		.		.		.		.
.		.		.		.		.		.
.		.		.		.		.		.
RN300	=	.59221	x	3791	=	2,245.06810	=	2,245	+ 1	= 2,246

* A five digit random decimal fraction is only useful on populations of 10,000 items or less. Additional random digits must be used in the decimal fraction if larger populations are sampled.

** Reference: The Rand Corporation, A Million Random Digits, Free Press, 1955, p. 355.

APPENDIX D
(Cont'd)

ATTACHMENT 4

GUIDELINES ON SAMPLE EXPANSION

The primary reason for continuing a sampling investigation is to determine if detected deficiencies are systematic or random, and aid in their evaluation. It may also be used as an aid in evaluating adverse trends and their root cause(s).

If the 95 percent upper-bound confidence limit (or probability interval) calculated from the sample is greater than 5 percent, the population of items is said to have failed the screen and further investigation is necessary. If one deficiency is detected in the initial minimum sample of items (or one or more different attributes in the case of VII.c) and no root cause can be identified, sample expansion in the entire population, including all attributes, will continue until it is determined that either the deficiency is a random occurrence of very low frequency, or a trend or programmatic deficiency is identified in the population (i.e., a potentially deficient stratum*). If a deficiency is detected in the initial sample and a root cause that implicates only a subset of attributes is identified, a reduced set of attributes will be considered in the sample expansion. If a sample is found to contain deficiencies or an adverse trend associated with items possessing a certain characteristic and not with items that do not possess this characteristic, a subset of items possessing these certain characteristics will be considered in the sample expansion. If deficiencies continue to be detected in the expanded sample, or two or more deficiencies of the same type are detected in the initial minimum sample, and they cannot be associated with a specific stratum, 100 percent of the population will be inspected or reviewed.

Sample expansion into a stratum will be required when it has been determined or hypothesized that the stratum contains and bounds the adverse trends or deficiencies of the type detected in the initial sample. Such a stratum may be identified by an adverse trend in the initial sample or by a root cause evaluation originating inside or outside the population of items being inspected or reviewed. Sample expansion into a stratum proceeds in one of the following ways:

- The stratum is identified completely, separated from the general population, items numbered sequentially from one to the total number and then randomly selected, or

---

* As used in the CPRT program, stratum will refer to either
1) a subset of items in the population,
2) a set of attributes of items in the population,
or 3) a set of attributes for a subset of items in the population.

APPENDIX D
(Cont'd)

ATTACHMENT 4
(Cont'd)

- The stratum is left within the general population and random sampling is continued in the general population until the required sample size is obtained in the stratum.

If there is no identified root cause for the initial deficiency or adverse trend, sample expansion in the general population is required to verify that the deficiency detected in the initial sample was a random occurrence of low frequency. When sample expansion is performed only for a stratum, the sample in the general population (minus the identified stratum) will be augmented with additional items to bring the general population sample back to the minimum 95/5 sample size.

For example, if one deficiency is found in an initial sample using the minimum 95/5 sample screen, and no stratum can be identified, then an additional 35 randomly selected items is needed (note: this is based on Bayes' theorem). If a deficiency is found in an initial minimum 95/5 sample, and a stratum is identified and removed for separate investigation, enough additional items must be randomly selected from the general population minus the identified stratum to bring the total sample back up to the minimum 95/5 sample size. In addition, the sample size in the stratum must total 95. Any items that were selected from the stratum in the initial sample are included as part of the sample expansion in the stratum. If no more deficiencies are detected, then the sample will pass the 95/5 screen and the conclusion will be made that the deficiencies are random and of very low frequency in the population (i.e., there was no programmatic breakdown).

Generating an expanded sample in the general population follows the same rules for generating the initial random sample. The sampling will start where the initial sample ended (see Table 3 of Attachment 3). Table 1 in Attachment 1 should be used as a guide for sample expansion in those cases where ISAP sample plans deviate from the general minimum 95/5 sample screen.

CERTIFICATE OF SERVICE

I, Robert K. Gad III, one of the attorneys for the Applicants herein, hereby certify that on January 31, 1986, I made service of the within "Applicants' Memorandum in Response to Board's Memorandum (Statistical Inferences from CPRT Sampling)," by mailing copies thereof, postage prepaid, to:

Peter B. Bloch, Esquire
Chairman
Administrative Judge
Atomic Safety and Licensing Board
U.S. Nuclear Regulatory Commission
Washington, D.C. 20555

Mr. James E. Cummins
Resident Inspector
Comanche Peak S.E.S.
c/o U.S. Nuclear Regulatory Commission
P.O. Box 38
Glen Rose, Texas 76043

Dr. Walter H. Jordan
Administrative Judge
881 W. Outer Drive
Oak Ridge, Tennessee 37830

Mr. William L. Clements
Docketing & Services Branch
U.S. Nuclear Regulatory Commission
Washington, D.C. 20555

Chairman
Atomic Safety and Licensing Appeal Panel
U.S. Nuclear Regulatory Commission
Washington, D.C. 20555

Chairman
Atomic Safety and Licensing Board Panel
U.S. Nuclear Regulatory Commission
Washington, D.C. 20555

Stuart A. Treby, Esquire
Office of the Executive Legal Director
U.S. Nuclear Regulatory Commission
7735 Old Georgetown Road
Room 10117
Bethesda, Maryland 20814

Mrs. Juanita Ellis
President, CASE
1426 S. Polk Street
Dallas, Texas 75224

Renea Hicks, Esquire
Assistant Attorney General
Environmental Protection Division
P.O. Box 12548, Capitol Station
Austin, Texas 78711

Ellen Ginsberg, Esquire
Atomic Safety and Licensing
Board Panel
U.S. Nuclear Regulatory Commission
Washington, D.C. 20555

Anthony Roisman, Esquire
Executive Director
Trial Lawyers for Public Justice
2000 P Street, N.W., Suite 611
Washington, D.C. 20036

Joseph Gallo, Esquire
Isham, Lincoln & Beale
1120 Connecticut Avenue, N.W.
Suite 840
Washington, D.C. 20036

Dr. Kenneth A. McCollom
Administrative Judge
Dean, Division of Engineering,
Architecture and Technology
Oklahoma State University
Stillwater, Oklahoma 74078

Mr. Lanny A. Sinkin
Christic Institute
1324 North Capitol Street
Washington, D.C. 20002

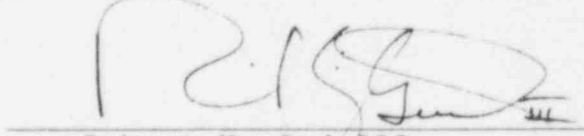
Ms. Billie Pirner Garde
Citizens Clinic Director
Government Accountability Project
1901 Que Street, N.W.
Washington, D.C. 20009

Mr. Robert D. Martin
Regional Administrator,
Region IV
U.S. Nuclear Regulatory Commission
Suite 1000
611 Ryan Plaza Drive
Arlington, Texas 76011

Elizabeth B. Johnson
Administrative Judge
Oak Ridge National Laboratory
P.O. Box X, Building 3500
Oak Ridge, Tennessee 37830

Geary S. Mizuno, Esquire
Office of the Executive
Legal Director
U.S. Nuclear Regulatory Commission
Maryland National Bank Bldg.
Room 10105
7735 Old Georgetown Road
Bethesda, Maryland 20814

Nancy Williams
Cygna Energy Services, Inc.
101 California Street
Suite 1000
San Francisco, California 94111



Robert K. Gad III