

Conf-911079--27

BNL-NUREG--46381

INCLUDING TEST ERRORS  
IN EVALUATING SURVEILLANCE TEST INTERVALS\*

DE92 007597

I.S. Kim, S. Martorell,\*\* W.E. Vesely,† and P.K. Samanta.

Engineering Technology Division  
Department of Nuclear Energy  
Brookhaven National Laboratory  
Upton, New York 11973, U.S.A.

RECEIVED  
FEB 10 1992

\*\*Universidad Politecnica of Valencia, Spain.

†Science Applications International Corporation, U.S.A.

FEB 10 1992

ABSTRACT

Technical Specifications require surveillance testing to assure that the standby systems important to safety will start and perform their intended functions in the event of plant abnormality. However, as evidenced by operating experience, the surveillance tests may adversely impact safety because of their undesirable side effects, such as initiation of plant transients during testing or wearing-out of safety systems due to testing.

This paper first defines the concerns, i.e., the potential adverse effects of surveillance testing, from a risk perspective. Then, we present a methodology to evaluate the risk impact of those adverse effects, focusing on two important kinds of adverse impacts of surveillance testing: (1) risk impact of test-caused trips and (2) risk impact of test-caused equipment wear. The quantitative risk methodology is demonstrated with several surveillance tests conducted at boiling water reactors, such as the tests of the main steam isolation valves, the turbine overspeed protection system, and the emergency diesel generators. We present the results of the risk-effectiveness evaluation of surveillance test intervals, which compares the adverse risk impact with the beneficial risk impact of testing from potential failure detection, along with insights from sensitivity studies.

I. BACKGROUND AND INTRODUCTION

Surveillance tests are required in nuclear power plants to detect failures in standby equipment as a means of assuring their availability in case of an accident. However, operating experience suggests that the tests may have adverse impact on safety,<sup>1,3</sup> that may be caused by test errors, e.g., human errors of omission or commission, including the potential for common-cause failures. This potential for adverse impact on safety is aggravated by the "overwhelming" amount of testing<sup>1,3</sup> presently required by Technical Specifications.

To address the problem of surveillance testing, i.e., the adverse effect on safety exacerbated by the significant amount of testing, the U.S. Nuclear Regulatory Commission (NRC) performed a series of studies. NUREG-1024<sup>2</sup> made recommendations to enhance the safety impact of surveillance requirements. NUREG-1366<sup>3</sup> implemented the recommendations by "qualitatively" examining all Technical Specifications surveillance requirements to identify those that should be improved. Four different types of adverse effects of testing were used in the NUREG-1366 study as screening criteria:

\*Work performed under the auspices of the U.S. Nuclear Regulatory Commission.

MASTER

810

(a) leading to a plant transient, (b) unnecessary wear to equipment, (c) unnecessary radiation exposure to plant personnel, and (d) unnecessary burden on plant personnel.

This paper summarizes the work<sup>4</sup> performed for the NRC by Brookhaven National Laboratory to help enhance the safety impact of surveillance testing. We first define the adverse effects of surveillance testing from a risk perspective, and then present a methodology to "quantify" the adverse risk impact, i.e., the penalty or increase in risk caused by the test. The quantitative methodology focuses on two important adverse effects, i.e., transients and equipment degradations. These adverse effects generate significant safety concerns because of: (1) plant abnormality which may challenge safety systems and plant operators and (2) equipment wear-out which increases the unavailability of the safety system or function, and thereby, reduces the plant's accident mitigating capability.

The risk impact of test-caused plant transients was evaluated by recognizing that the transients, which cause or require a reactor scram, are initiating events as typically called in probabilistic risk assessments (PRAs). The risk impact of equipment degradations was assessed using a test-caused component degradation model which was developed in this study from considerations of the stresses on equipment and degradation mechanisms induced by testing and aging.

The methods for evaluating the adverse effects of testing were applied to several surveillance tests at boiling water reactors, such as those on the main steam isolation valves, the turbine overspeed protection system, and the emergency diesel generators. The risk associated with these tests was assessed using a PRA conducted in the NUREG-1150 study.<sup>5</sup> Risk-effectiveness was evaluated by comparing the risk impact of plant transients and equipment degradations from these tests to the beneficial risk impact of testing resulting from the detection of failures. Sensitivity studies were also carried out on risk impact versus test interval.

Section 2 of this paper defines the adverse effects of testing from a risk perspective. Sections 3 and 4 present the methodology to evaluate the risk impacts of testing associated with transients and equipment degradations, along with the results of the risk-effectiveness evaluation and sensitivity studies. Section 5 gives our conclusions.

## 2. RISKS ASSOCIATED WITH A TEST

Surveillance testing may have two different types of risk impact on the plant: a beneficial impact, i.e., the reduction in risk, and an adverse impact, i.e., the increase in risk. The beneficial risk contribution results from the detection of failures which occur between tests. This risk contribution "detected" by a test here will be called  $R_D$ . The adverse risk contribution results from degradations or failures that are due to or related to the test, and from the component unavailability during or as a result of the test. This contribution "caused" by a test will be called  $R_C$ .

The test-detected risk contribution,  $R_D$ , resulting from the detection of failures can be quantified in the framework of a PRA, as demonstrated in reference 6. The test-caused contribution,  $R_C$ , may have several different kinds of contributions. Table 1 lists the different risk contributions which can be associated with a test, along with the root causes of the risk.

From Table 1, the test-caused risk can be expressed in a general form as

$$R_C = R_{\text{trip}} + R_{\text{wear}} + R_{\text{state}} + R_{\text{down}} \quad (1)$$

where, for any specific test, some contributions may be irrelevant or insignificant compared to the others. When a test program or procedure is evaluated for its risk-effectiveness by conducting tests on a number

Table 1. Test-Caused Risk Contributions and Their Root Causes

Identifier	Risk Contribution from Test	Root Causes of the Risk
$R_{trip}$	Risk from trips	Human error, equipment failure, procedure inadequacy.
$R_{wear}$	Risk from equipment wear	Inherent characteristics of the test, procedure inadequacy, human error.
$R_{state}$	Risk from misconfigurations or errors in component restoration	Human error, procedure inadequacy.
$R_{down}$	Risk associated with downtime in carrying out the test	Unavailability of the component during the test. Affected by the test override capability.

of individual components, then the contributions for each test plus the contributions from any test interactions need to be considered.

In addition to those effects defined in Table 1, two other adverse effects may be sometimes encountered: radiation exposure to plant personnel and unnecessary burden of work on plant personnel. These adverse effects differ from those in the table in that: (1) the radiation exposure to plant personnel is not amenable to a risk analysis based on the core-damage frequency as a risk measure, and (2) the unnecessary burden of work on plant personnel in general is not subject to a risk analysis. Although excluded from the quantitative risk analysis, these adverse effects can be considered qualitatively along with the results of quantitative risk analysis for the evaluation of surveillance requirements.

Among the various root causes of the risk delineated in Table 1, human errors of component restoration following tests are the root cause which previous studies<sup>7,8</sup> concentrated on to address adverse effects of testing. In terms of the risk contributions in the table, the studies focused on  $R_{state}$  that is most likely to be caused by human errors, with some consideration of  $R_{down}$ .

The risk also can be evaluated for a specific root-cause, such as human errors (or more specifically, errors of omission or commission). For instance, presume that human errors during a test may cause a transient and also may cause the components not to be restored to the normal status. The risk contribution due to potential human errors during the test can then be estimated by first evaluating the contributions of the risk from trips,  $R_{trip}$ , and from component restoration error,  $R_{state}$ , due to only human errors, and then adding the contributions.

In evaluating the risk-effectiveness of a test (or group of tests), the test-detected contribution,  $R_D$ , can either be compared to specific test-caused contributions or to all relevant contributions constituting  $R_C$ . When  $R_C$  is assessed by considering all significant contributions, we can define the risk-effectiveness of a test as follows: a test is risk-effective if  $R_D > R_C$ , otherwise it is risk-ineffective. If only specific contributions are considered, then the evaluation of the risk-effectiveness of the test is considered with regard to the specific test-caused contributors. For example, if test-caused risk contributions due to trips,  $R_{trip}$ , are only considered and we assess that  $R_D > R_{trip}$ , then we can say that the test is risk-effective with regard to test-caused trips. When more test-caused contributions are considered, then broader conclusions can be reached.

### 3. RISK IMPACT OF TEST-CAUSED TRANSIENTS

The operating history of nuclear power plants suggests that a surveillance test conducted at power may cause a transient that will lead to or require a reactor trip. The risk impact of such transients depends on the various responses of the plant safety systems, and also on plant operators. Typically, in PRAs the various plant or operator responses that may affect the plant risk are taken into account using event trees to delineate the progressions of accident sequences and system fault trees to identify the failure modes and their effects on the system unavailabilities. Hence, the risk contribution from test-caused transients to the plant risk can be evaluated within the framework of a PRA model.

The risk impact of a test-caused transient,  $R_{\text{trip}}$ , can be evaluated through that of the PRA initiating event group associated with the transient:

$$R_{\text{trip}} = \phi R_{\text{IE}j} \quad (2)$$

where  $R_{\text{IE}j}$  denotes the risk impact of the  $j$ -th initiating event group which is assumed to be associated with the test-caused transient, and  $\phi$  is the proportion by which the frequency of the PRA initiating event group is attributable to these transients. The proportion,  $\phi$ , can be estimated by analyzing plant operating data:

$$\phi = \frac{N_{\text{test}}}{N_{\text{IE}j}} \quad (3)$$

where,

$N_{\text{test}}$  = the number of test-caused transient events, and

$N_{\text{IE}j}$  = the number of transient events belonging to the initiating event group associated with the test-caused transient.

To obtain  $\phi$ , the test-caused transients must be associated with the relevant initiating event groups. For this purpose, the EPRI (Electric Power Research Institute) transient categories<sup>9</sup> that were originally developed to analyze the historical transient events in a study of anticipated transients without scram (ATWS) can be used. Such use will facilitate and improve the accuracy of the data analysis, because the extent of detail on the test-caused transients and the PRA initiating event groups is usually quite different. The ATWS study defined 37 transient categories for BWRs based on the different characteristics of a variety of transient events that had occurred or might occur in the plants.

For sensitivity studies (in terms of risk impact versus test interval that will be discussed later), we first get the following equation for the probability,  $p_{\text{trip}}$ , that a transient will occur during or as a result of a test:<sup>4</sup>

$$p_{\text{trip}} = I_j T \phi \quad (4)$$

where  $T$  and  $I_j$  denote the test interval and the frequency of the  $j$ -th initiating event group used in the PRA model, respectively. Substituting an expression for  $\phi$  from Equation (3) into Equation (2) we have:

$$R_{\text{trip}} = \frac{P_{\text{trip}}}{I_j T} R_{\text{IE-j}} \quad (5)$$

These formulas discussed above were used in the framework of a NUREG-1150 PRA for a BWR to evaluate the risk impact and effectiveness of the following tests: a) a quarterly test of the main steam isolation valve (MSIV) operability and b) a weekly test of the turbine overspeed protection system (TOPS).

Table 2 shows the BWR transient categories that were identified as being associated with the tests, based on the test characteristics and the effects of the test-caused transients on the plant. For example, a TOPS test may cause the turbine control valve to fail closed, resulting in high steam pressure in the main steam system, and consequently, in a turbine trip. Hence, the transient due to the TOPS test can be classified into Category 3, "Turbine trip," and Category 13, "Turbine bypass or control valves cause increased pressure (closed)."

Table 2. Association of Transient Categories Relevant to Test-Caused Transients with PRA Initiating Event Groups

Test	Transient Category	Description	PRA Initiating Event Group
MSIV	6	Inadvertent closure of one MSIV	T2
	7	Partial MSIV closure	T2
TOPS	3	Turbine trip	T3A
	13	Turbine bypass or control valves cause increased pressure (closed)	T3A

The transient categories were then associated with the initiating event groups modeled in the plant-specific PRA, based on the characteristics of the categories and the groups. The plant-specific PRA initiating event groups which were identified to be associated with the transient categories are listed in Table 2. Categories 3 and 13 of the TOPS test are associated with initiating event group T3A, i.e., transients with the power conversion system initially available except those due to an inadvertent open relief valve in the primary system and those involving loss of feedwater. Categories 6 and 7 of the MSIV operability test are associated with initiating event group T2 which incorporates transients with the power conversion system unavailable.

The results of the core-damage frequency impact of test-caused transients,  $R_{\text{trip}}$ , and the probability that a transient will occur during or as a result of a test,  $P_{\text{trip}}$ , are the following:

a) For quarterly MSIV test:

$$R_{\text{trip}} = 1.8\text{E-}7 \text{ per reactor year}$$

$$P_{\text{trip}} = 6.7\text{E-}2 \text{ per test}$$

b) For weekly TOPS test:

$$R_{\text{trip}} = 3.7\text{E-}8 \text{ per reactor year}$$
$$p_{\text{trip}} = 1.7\text{E-}3 \text{ per test}$$

Comparing the results for the two different tests, we see that, for these test intervals, the risk contribution from test-caused transients and the probability of a transient occurring during a test for the MSIV test is greater than those for the TOPS test by a factor of 5 and 4, respectively.

We can also examine whether the test is risk-effective with respect to test-caused transients by comparing the value of  $R_D$  to that of  $R_{\text{trip}}$ . The quarterly MSIV test is risk-effective, because the  $R_D$  is  $5.2\text{E-}7$  per year which is larger than the  $R_{\text{trip}}$ ; the risk-effective margin is  $3.4\text{E-}7$  per year. The risk-effectiveness of the TOPS test could not be evaluated based on the NUREG-1150 PRA, because the turbine control valves are not modeled in the PRA. Hence, for the TOPS test, only the quantitative values of  $R_{\text{trip}}$  and  $p_{\text{trip}}$  can be taken into account in the evaluation of the test, unless the value of  $R_D$  is obtained by modifying the PRA model.

Figure 1 shows the result of the sensitivity study of MSIV operability testing for three different kinds of core-damage frequency impacts to the variation of the test interval,  $T$ : (1) the test-caused core-damage frequency contribution due to transients,  $R_{\text{trip}}$ , (2) the test-detected core-damage frequency contribution,  $R_D$ , and (3) the total core-damage frequency impact of the test,  $R_T$ , which is simply the sum of  $R_{\text{trip}}$  and  $R_D$ .  $R_{\text{trip}}$  decreases as  $T$  is increased, because less transients are expected as the test is conducted less frequently. However,  $R_D$  increases with the increasing test interval, because the test is more likely to detect a failure.

The risk-effectiveness of the test on test-caused transients also can be seen by comparing the test-detected risk contribution to the test-caused risk contribution due to transients. In the region where  $T > 54$  days,  $R_D$  is larger than  $R_{\text{trip}}$ , and therefore, the test is risk-effective. In the other region where  $T \leq 54$  days, the test is risk-ineffective.

An important conclusion relevant to the redefinition of a standard test interval is that the interval for MSIV operability testing, i.e., 91 days, can be extended without undue increase in the risk impact. For example, if the test interval is extended to 150 days,  $R_D$  increases because the test is more likely to detect failures, while  $R_{\text{trip}}$  decreases because less testing during a given time period will result in less transients. However, as shown by a dotted curve in Figure 1, the total risk impact of the test,  $R_T$ , only marginally increases when  $T$  is changed from 91 days to 150 days. ( $R_T$  increases from  $6.99\text{E-}7$  per year to  $9.64\text{E-}7$  per year.)

In this study, we used the LER data base for 30 BWRs for 1985, assuming that the operability of MSIVs is tested quarterly at all the plants. However, the data analysis indicated that some plants test the operability of MSIVs more frequently; e.g., the operators of a nuclear plant were performing a biweekly surveillance when the test failure occurred in the plant. If we assume that the minimum test interval of 54 days is also applicable to this plant, we can say that the biweekly test is risk-ineffective with regard to test-caused transients, because the interval is shorter than 54 days. Even if we consider other types of adverse risk impacts and they are not negligible compared to  $R_{\text{trip}}$ , the test will be risk-ineffective.

Sensitivity analyses, such as that shown in Figure 1, can be very useful in defining test intervals. However, they should be carefully interpreted. In Figure 1, the sensitivity curves of  $R_{\text{trip}}$  and  $R_T$  to the variation of  $T$  are based on the assumption that the probability,  $p_{\text{trip}}$ , of a transient occurring during testing is constant. However, the value of  $p_{\text{trip}}$  may change (tend to increase), especially when the test

is conducted far less frequently than it used to be, because the operators are more likely to make errors. Therefore, when considering an extension of test interval based on the sensitivity analyses, one should not prolong the test interval too much, e.g., by more than a factor of two. The extension of the test interval mainly depends on the likelihood that  $p_{trip}$  will vary following the change of the test frequency.

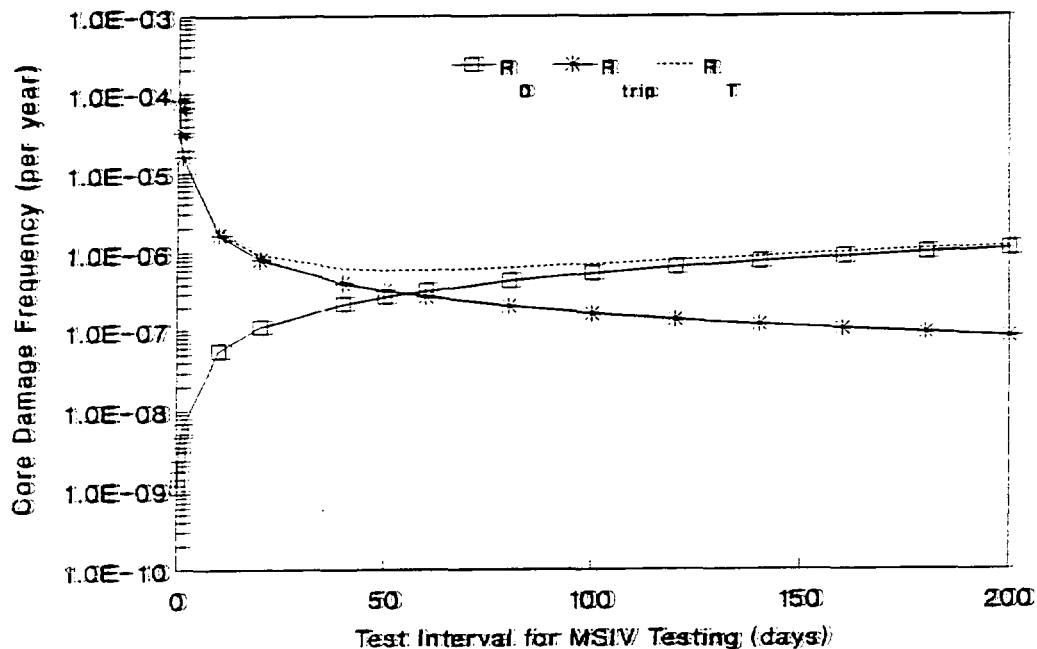


Figure 1. Sensitivity of the core-damage frequency impact to the test interval for the main steam insolation valve testing ( $R_D$  = test-detected risk impact;  $R_{trip}$  = test-caused risk impact due to transients;  $R_T$  = total risk impact of the test).

#### 4. RISK IMPACT OF TEST-CAUSED DEGRADATIONS

The safety-significant components of nuclear power plants, such as a diesel generator or an auxiliary feedwater pump, are tested so often—generally monthly and sometimes more often—that the tests may lead to progressive wear-out of the equipment due to the accumulation of the test-caused degradation effects. Furthermore, as time passes the component will also show aging effects, such as corrosion or erosion. The accumulating test-caused degradation and aging effects will increase the unavailability of the component, and thereby, the unavailability of the associated safety system and function. This increase will, in turn, reduce the plant's accident mitigating capability.

The degradation from testing and aging effects are induced by two kinds of stresses, i.e., demand and standby stresses. Demand stress acts on equipment only when the equipment is asked to function or is operating. Standby stress acts while it is in the standby state. For standby components which are periodically tested, generally the combination of both stresses causes the equipment to degrade, and ultimately, to fail.

Based on the concept of stress on equipment and the characteristics of the degradation mechanisms caused by testing and aging, we can formulate the following test-caused component degradation model:

$$q(n,t) = \rho(n) + \int_{nT}^{nT+t} \lambda(n,t') dt' \quad \text{for } t \in [0, T] \quad (6)$$

$$\rho(n) = \rho_0 + \rho_1 n \quad (7)$$

$$\lambda(n,t) = \lambda_0 + \lambda_1 n + \alpha u \quad \text{for } u \in [0, nT+t] \quad (8)$$

where,

$n$  = number of tests performed on the equipment

$t$  = time elapsed since the last test

$\rho(n)$  = failure probability for demand caused failures

$\lambda(n,t)$  = standby failure rate (per unit time) for failures occurring between tests

$T$  = test interval

$nT + t$  = time since the last renewal point

$\rho_0$  = residual demand-failure probability

$\rho_1$  = test degradation factor associated with demand failures

$\lambda_1$  = test degradation factor for standby time-related failures

$\alpha$  = aging factor associated with pure aging

Equations (6) to (8) represent a model which has been linearized from the original, non-linear test-caused degradation model.<sup>4</sup> This linear model can be used for most purposes and is used in this paper.

In Equation (6) the unavailability,  $q(n,t)$ , and the standby failure rate,  $\lambda(n,t)$ , are represented as a function of  $n$  and  $t$ . The reason for this functional notation is that the standby failure rate is assumed to be affected by not only the standby time, but also by test-caused degradation. Therefore, component unavailability becomes a function of the number of tests performed on the component since the last renewal point, as well as the time elapsed since the last renewal. However, the demand failure probability is represented in Equation (6) as a function of only the number of tests,  $n$ , i.e., it is assumed that the demand-failure probability depends only on how many tests have been conducted on the component.

The expressions for the two basic degradation parameters,  $\rho(n)$  and  $\lambda(n,t)$ , are formulated in Equations (7) and (8) in terms of their variables  $n$  and  $t$ . In Equation (8) the time-dependent aging mechanism on the standby failure rate is represented by a Weibull distribution.

The test-caused component degradation model, Equations (6) to (8), provides a means to estimate the time-dependent component unavailability and its resultant risk impact as a function of the number of tests on the component and the time elapsed since the last overhaul time.



Let

$\bar{R}_{C;n}$  = the average increase in core-damage frequency or test-caused risk contribution resulting from test-caused degradations of  $n$  tests on the equipment

We can evaluate  $\bar{R}_{C;n}$  using the following formula:

$$\begin{aligned} \bar{R}_{C;n} &= \text{the average risk level between } [t_0, t_0 + T] - \text{the average risk level between } [0, T] \\ &= \Delta \bar{q}_n [R_1 - R_0] \\ &= (p_1 \rho_0 n + \frac{1}{2} p_2 \lambda_0 T n) [R_1 - R_0] \end{aligned} \quad (9)$$

where  $\Delta \bar{q}_n$  denotes the average increase in component unavailability that results from  $n$  tests, and only the test-caused degradation effect is taken into account without considering the aging effect, i.e.,  $\alpha = 0$ .

Based on these formulas, we can establish the following criterion on the number of tests for risk-effectiveness with regard to test-caused degradation:

$$n < \frac{\frac{1}{2} \lambda_0 T}{\rho_0 p_1 + \frac{1}{2} \lambda_0 p_2 T} \quad \therefore \quad \begin{array}{l} n\text{-th test risk-effective with regard} \\ \text{to test-caused degradation} \end{array} \quad (10)$$

For the  $n$ -th test to be risk-effective with regard to test-caused degradation, the number of tests performed on the component since the last overhaul should satisfy the above criterion. When the number of tests on the component is less than the value of the right-hand side in the criterion, then the contribution to core-damage frequency caused by the test will be less than the contribution to core-damage frequency detected by the test, and vice versa.

The test-caused component degradation model not only incorporates aging effects, but separately takes into account test-caused degradation. However, the degradation model and the formulas for evaluating the risk impact associated with such degradations are based on the following assumptions that may shed light on some limitations in the use of the approaches:

- (1) Test-caused component degradations affect demand failure probability, and also standby failure rate; i.e., the component will be more vulnerable to both demand and standby time-related failures as more tests are performed on the component.
- (2) The standby time-related failure rate increases because of test-caused degradation effects, as well as aging effects. However, aging does not affect the probability that the component will fail upon demand, i.e., the demand-failure probability.
- (3) The time-dependent aging mechanism on the standby failure rate can be represented by a Weibull distribution.
- (4) The demand degradation or failure mechanism is not affected by time. In other words, the demand failure probability depends on only the number of tests performed on the equipment, but not on the idle or dormant time.

Using the test-caused degradation model we have discussed, a sensitivity study was performed on the risk impact of test-caused equipment degradations versus test interval. The risk impact was evaluated in the framework of a NUREG-1150 PRA model. We chose the emergency diesel generator as the sample component, because of the concern about test-caused degradations on this component and the availability of the necessary reliability data to estimate the degradation parameters of the model. However, the method presented here can also be applied to any other component.

The values of the degradation parameters, such as  $p_1$  and  $p_2$ , were estimated for diesel generators under the following assumption:

When the number of tests is large, the average increase in component unavailability which is evaluated by the test-caused component degradation model is the same as that estimated by the aging model.<sup>10</sup>

Figures 2 and 3 show the results of the sensitivity study for monthly and quarterly testing of the diesel generators, respectively. They show the sensitivity of three different kinds of core-damage frequency impacts to the variation in the number of tests: (1) the test-detected core-damage frequency contribution,  $\bar{R}_D$ , (2) the test-caused core-damage frequency contribution due to equipment wear,  $\bar{R}_{C,m}$ , and (3) the total core-damage frequency impact of the test,  $\bar{R}_{T,m}$ .

For monthly testing, the component degradation model indicates that the test is risk-effective until 61 tests have been performed, i.e., approximately 5 years after the last overhaul. For quarterly testing, the model indicates that the test is risk-effective until 111 tests have been performed, i.e., about 28 years after the last renewal time. However, when the test is no longer risk-effective, the total risk impact for quarterly testing is greater than that for monthly testing by approximately a factor of 3.

The numerical results from this analysis should be interpreted cautiously, because the values of the degradation parameters, which are component-specific, were estimated using the results from reliability studies of a number of different diesel generators. For more meaningful results, the model should be used with the degradation parameters for the specific diesel generator whose risk-effectiveness is to be evaluated.

## 5. CONCLUSIONS

The safety significance and risk-effectiveness of surveillance test requirements can be evaluated by explicitly considering the adverse effects of testing, based on the concepts and methods discussed in this paper. The results of quantitative risk evaluation can be used in the decision-making process to establish the safety significance of the surveillance testing and to screen surveillance requirements. These results should be used in conjunction with qualitative evaluations from engineering considerations and operating experience, such as qualitative evaluations of radiation exposure to plant personnel from the tests and test-caused operator burden.

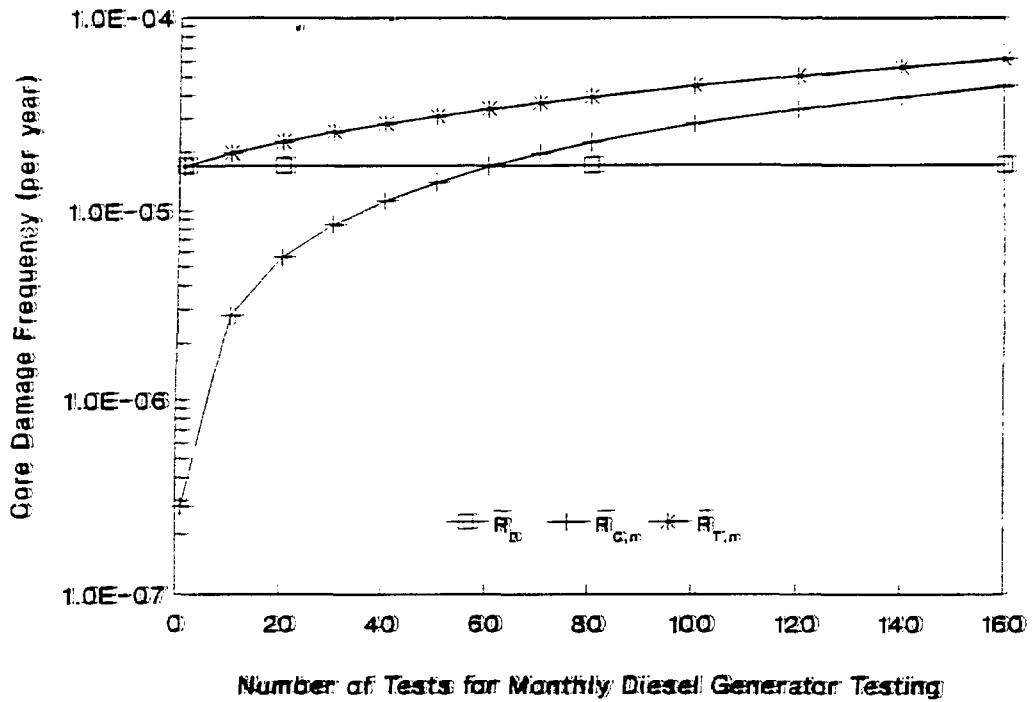


Figure 2. Evaluation of risk-effectiveness for monthly diesel generator testing ( $\bar{R}_D$  = test-detected risk impact;  $\bar{R}_{C,n}$  = test-caused risk impact due to equipment wear;  $\bar{R}_{T,n}$  = total risk impact of the test)

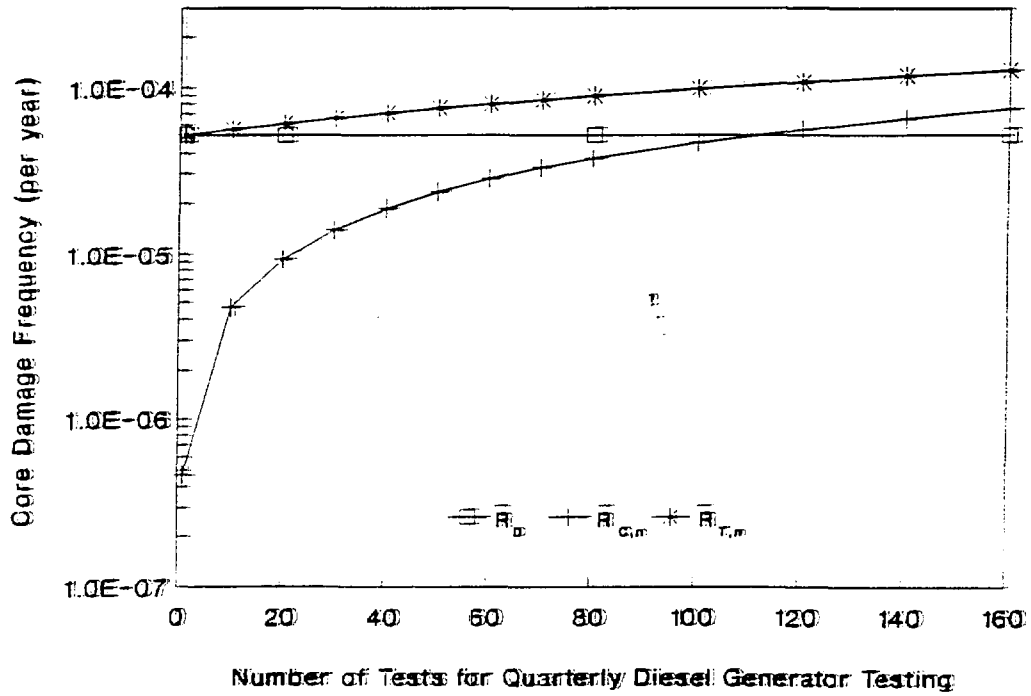


Figure 3. Evaluation of risk-effectiveness for quarterly diesel generator testing ( $\bar{R}_D$  = test-detected risk impact;  $\bar{R}_{C,n}$  = test-caused risk impact due to equipment wear;  $\bar{R}_{T,n}$  = total risk impact of the test)

## REFERENCES

1. U.S. Nuclear Regulatory Commission (NRC), "A Survey by Senior NRC Management to Obtain Viewpoints on the Safety Impact of Regulatory Activities from Representative Utilities Operating and Constructing Nuclear Power Plant," NUREG-0839, August 1981.
2. U.S. NRC, "Technical Specifications -- Enhancing the Safety Impact," NUREG-1024, November 1983.
3. R. Lobel and T.R. Tjader, "Improvements to Technical Specifications Surveillance Requirements," NUREG-1366, August 1990.
4. I.S. Kim, S. Martorell, W.E. Vesely, and P.K. Samanta, "Quantitative Evaluation of Surveillance Test Intervals Including Test-Caused Risks," NUREG/CR-5775, BNL-NUREG-52296, in press.
5. U.S. NRC, "Severe Accident Risks: An Assessment for Five U.S. Nuclear Power Plants," NUREG-1150, Vols. 1-2, Final Report, December 1990.
6. P.K. Samanta, S.M. Wong, and J. Carbonaro, "Evaluation of Risks Associated with AOT and STI Requirements at the ANO-1 Nuclear Power Plant", NUREG/CR-5200, BNL-NUREG-52024, August 1988.
7. G.E. Apostolakis and P.P. Bansal, "Effect of Human Error on the Availability of Periodically Inspected Redundant Systems," IEEE Transactions on Reliability, Vol. R-26, No. 3, August 1977.
8. T.P. McWilliams and H.F. Martz, "Human Error Considerations in Determining the Optimum Test Interval for Periodically Inspected Standby," IEEE Transactions on Reliability, Vol. R-29, No. 4, October 1980.
9. A.S. McClymont and B.W. Poehlman, "ATWS: A Reappraisal, Part 3: Frequency of Anticipated Transients," EPRI NP-2230, January 1982.
10. W.E. Vesely, R.E. Kurth, and S.M. Scalzo, "Evaluation of Core Melt Frequency Effects due to Component Aging and Maintenance", NUREG/CR-5510, June 1990.