

October 12, 2001

The Honorable Richard A. Meserve  
Chairman  
U.S. Nuclear Regulatory Commission  
Washington, DC 20555-0001

SUBJECT: THE REVISED REACTOR OVERSIGHT PROCESS

Dear Chairman Meserve:

During our 485<sup>th</sup> meeting on September 5-7, 2001, the Advisory Committee on Reactor Safeguards met with representatives of the NRC staff to discuss the revised Reactor Oversight Process (ROP). We continued our deliberations during our 486<sup>th</sup> meeting on October 4-6, 2001. This matter was also discussed during meetings of the ACRS Plant Operations Subcommittee on December 6, 2000, May 9, 2001, and July 9, 2001. In addition, the ACRS Subcommittees on Plant Operations and Fire Protection held meetings with licensees on June 13, 2000, and June 27, 2001, and held meetings with Regions III and IV on June 14, 2000, and June 28, 2001, respectively. During our review, we had the benefit of the documents referenced.

## **BACKGROUND**

The ROP utilizes the results of performance indicators (PIs) and baseline inspection findings to determine the appropriate regulatory action to be taken in response to a licensee's performance. The escalation of the regulatory responses is specified in the action matrix, which the staff developed as part of the ROP. This ROP has been in effect for nearly all licensees for about one year. The staff has conducted an assessment of the state of the ROP and recognizes that it is still a process in development.

The ACRS has previously commented on various aspects of the ROP and provided recommendations to the staff regarding potential process improvements. We remain convinced that the ROP is more objective and understandable than the former oversight process and represents a significant improvement. This report discusses some specific questions that the Commission raised to the ACRS, and offers some additional thoughts on potential improvements in the ROP.

In the Staff Requirements Memorandum dated April 5, 2000, the Commission requested the ACRS to:

- (1) Review the use of PIs in the ROP to ensure that the PIs provide meaningful insight into aspects of plant operation that are important to safety.

- (2) Review the initial implementation of the significance determination processes (SDPs), and assess the technical adequacy of the SDP to contribute to the ROP.

The current PIs do provide meaningful insight into plant performance. However, there is a need to redefine the thresholds for some of the PIs to provide better input to the ROP. In particular, the numerical values for the white/yellow and yellow/red thresholds for the initiating event and mitigation system PIs are not useful and should be revised. The color bands for the PIs and SDPs associated with all the cornerstones have similar implications with respect to agency action and, therefore, the thresholds should be commensurate with their respective safety significance.

The most immediate and pressing need for the ROP is to improve the SDP tools. Some SDPs are incomplete and, in cases such as fire protection, overly subjective. The technical adequacy of the risk-based SDPs depends on the availability and quality of a relevant probabilistic risk assessment (PRA). Thus, the SDP for at-power situations provides meaningful risk information. For routine findings that are predominantly of very low, low, and moderate safety significance, the process is probably adequate. The threshold values for the risk-based SDPs are appropriate.

We continue to believe that a documented review of the SDP worksheets and SPAR models (as well as the underlying SAPHIRE computer code) is essential to public confidence in the ROP.

An SDP based on low-power and shutdown PRAs or other shutdown management tools is needed to characterize findings during these modes of operation. In addition, the fire protection SDP involves very qualitative inputs to a quantification process of uncertain pedigree. This SDP is probably useful for its intended purpose, however, it may be hard to defend and justify to the public. Even though this SDP calculates the change in core damage frequency (CDF), the SDP is really intended to provide an indication of the degradation of defense in depth for fire protection as defined in 10 CFR Part 50, Appendix R.

Presently, concurrent performance deficiencies are assessed collectively, as applicable, to determine the total change in CDF, but each performance deficiency is assigned a color individually. There may be instances in which conclusions could be altered if the results are considered collectively, and thus such collective results should be considered in the action matrix.

## **DISCUSSION**

An important premise of the ROP is that there should be a graded regulatory response to inspection findings and PI results. Although a graded response to oversight findings is a desirable attribute, the inputs to the action matrix that implements this response must be produced in a way that justifies the resulting response. This is especially true for the right-hand columns of the matrix which could lead to severe regulatory responses.

The current ROP uses different technical bases to establish the thresholds for the PIs and inspection findings. In particular:

- On the basis of its review of recent operating history, the staff set the green/white thresholds for the PIs for initiating events and mitigating systems at the 95<sup>th</sup> percentile of peer performance for the given PI. By contrast, the staff based the white/yellow and yellow/red thresholds on an assessment of the value of a PI corresponding to increases in CDF of  $10^{-5}$  and  $10^{-4}$  per reactor year, respectively.
- The staff set the PI thresholds for barriers, emergency preparedness, occupational radiation safety, public radiation safety, and physical protection by considering technical specification limits, the number of noncompliances with regulatory requirements, and other absolute measures.
- The staff based the green/white, white/yellow, and yellow/red thresholds for SDP results on increases in CDF of  $10^{-6}$ ,  $10^{-5}$ , and  $10^{-4}$  per reactor year, respectively. This is true for the initiating event, mitigating system, and fire protection cornerstones. The other SDPs do not have a PRA basis and take a deterministic and defense-in-depth approach to establish thresholds for safety significant issues.

These different bases for defining the various thresholds raise questions regarding the kinds of information that the PIs and SDPs provide and the consistency of the meaning of the thresholds across the PIs and SDPs. These different thresholds are based on expert judgment that the degradation in performance associated with each color band is appropriately linked to a corresponding regulatory response<sup>1</sup>.

It is from this viewpoint that we believe it is necessary to reconsider the definitions of the white/yellow and yellow/red thresholds for initiating events and mitigating systems, which as we noted above were based on an attempt to assess the value of a PI corresponding to increases in CDF.

We have noted previously that it is difficult to generically assess the risk impact of changes in a PI. The associated changes in risk tend to depend strongly on plant-specific features. This approach, however, has a deeper, more intractable flaw. Specifically, it focuses on the change in CDF that results from changes in a single, isolated parameter assuming that all other factors that can affect CDF remain constant. A realistic assessment of the change in CDF cannot be related to the change in a single PI. Thus, in some cases, the use of this approach to select white/yellow and yellow/red thresholds has led to values for these thresholds that, in our judgment and that of many of the staff and the industry, are too high to be meaningful. Regulatory attention would increase at much lower levels.

The white/yellow and yellow/red thresholds for the PIs for initiating events and mitigating systems should be set in terms of an expert judgment of what values should in fact trigger the regulatory response associated with the threshold. Although general considerations for the selection of thresholds for PIs and SDPs are discussed in SECY-99-007, the expert judgment

---

<sup>1</sup> The color bands for the ROP are called “constructed scales” in decision analysis. Ensuring the consistency of the bands of these scales is what decision analysts commonly call “performing sanity checks,” and such checks are among the most important steps in a decisionmaking process. In our report on the NRC Safety Research Program (NUREG-1635, Vol. 4), we recommended that the staff initiate a program of research to investigate how best to use formal decisionmaking methods in regulatory decisions.

process that the staff used to develop the initial values for the thresholds for the non risk-based PIs and SDPs and the corresponding equivalency of the combination of findings in the action matrix have not been well documented. The NRC has been a pioneer in the use of scrutable expert judgment processes, and it is unfortunate that the use of expert judgment in a process as central to the NRC's mission as the ROP lacks the traceability of other NRC uses of expert judgment. Formal decision analysis could be helpful in making the selection of thresholds and the action matrix more objective and scrutable.

In assessing the need to revise the current PIs and develop new PIs, we believe that the staff responsible for the ROP should consider the work being done in other parts of the agency. For example, the review of operating experience for the reactor core isolation cooling (RCIC) system for BWRs (NUREG/CR-5500, Vol. 7) shows that the dominant failure modes involve system failures while running and human failures to recover the system (i.e., failures that are not part of the unavailability calculations that the ROP requires). In analyzing the operating experience, the analysts distinguished between two contexts of RCIC system operation: (1) short-term missions (less than 15 minutes), in which the system must inject water into the reactor vessel following a scram with feedwater available and the main isolation valves open, and (2) long-term missions, in which the system must inject water into the reactor vessel following a scram with feedwater unavailable and/or the reactor vessel isolated. The average system unreliability in these two contexts differs by a factor of 2. The ROP green/white threshold for RCIC system unavailability is 0.04 and makes no distinction between the two contexts identified in the study driven by operating experience. Since unreliability is a metric that includes all potential failure modes, it should be included in the PIs.

We continue to believe that it is important that there be consistency in the definition of terms like "unavailability" which are used in the PIs. Inconsistencies in technical terms that the agency uses in several major activities make comparisons and communication, both internally and externally, difficult.

The ROP is an evolving process. The staff has done an excellent job establishing the basic framework in a relatively short period of time considering the scope of this project. We look forward to continued interactions with the staff on this very important matter.

Additional comments by ACRS Members George E. Apostolakis, Thomas S. Kress, and Steven L. Rosen are presented below.

Sincerely,

/RA/

George E. Apostolakis  
Chairman

#### **References:**

1. Staff Requirements Memorandum dated April 5, 2000, from Annette L. Vietti-Cook, Secretary, NRC, to Dr. John T. Larkins, Executive Director, ACRS, Subject: Staff

- Requirements - Meeting on March 2, 2000, with ACRS on Risk Informing 10 CFR Part 50.
2. Letter dated March 15, 2000, from Dana A. Powers, ACRS Chairman, to Richard A. Meserve, Chairman, NRC, Subject: Revised Reactor Oversight Process.
  3. NRC Inspection Manual, Manual Chapter 0609, Appendix A, Significance Determination of Reactor Inspection Findings for At-Power Situations, February 5, 2001.
  4. NUREG/CR-5500, Vol. 7, Reliability Study: Reactor Core Isolation Cooling System, 1987 - 1993, Idaho National Engineering and Environmental Laboratory, September 1999.
  5. U. S. Nuclear Regulatory Commission, SECY-99-007, Recommendations for Reactor Oversight Process Improvements, January 8, 1999.
  6. U. S. Nuclear Regulatory Commission, SECY-99-007A, Recommendations for Reactor Oversight Process Improvements (Follow-up to SECY 99-007), March 22, 1999.
  7. U. S. Nuclear Regulatory Commission, SECY 01-0114, Results of the Initial Implementation of the New Reactor Oversight Process, June 25, 2001.
  8. U. S. Nuclear Regulatory Commission, SECY 00-0049, Results of the Revised Reactor Oversight Process Pilot Program, February 24, 2000.
  9. U. S. Nuclear Regulatory Commission Inspection Manual, Manual Chapter 0305, Operating Reactor Assessment Program, March 23, 2001.
  10. NRC Inspection Manual, Manual Chapter 0609, Significance Determination Process, February 27, 2001.
  11. Advisory Committee on Reactor Safeguards, NUREG-1635, Vol 4, Review and Evaluation of the Nuclear Regulatory Commission Safety Research Program, A Report to the U.S. Nuclear Regulatory Commission, May 2001.

**ADDITIONAL COMMENTS BY ACRS MEMBERS  
GEORGE E. APOSTOLAKIS, THOMAS S. KRESS, AND STEPHEN ROSEN**

We agree with the recommendations and comments of our colleagues. The intent of our comments is to elaborate on the expert judgment process.

In any decisionmaking situation, the most important requirement is that the decisionmaker's judgments be consistent. This is particularly important for the ROP because the bases for the inputs to the action matrix are different.

One of the columns of the action matrix treats two white inputs and one yellow input (for one degraded cornerstone) as being equivalent. This means that the staff's judgment is that two white inputs signify a certain degradation in performance which is about the same as that corresponding to one yellow finding in the sense that the resulting regulatory response should be the same. For consistency in defining these color bands, one would have to address questions such as the following:

- Does the yellow band for the initiating event PI indicate a degradation in performance that is similar to that indicated by the yellow band for a mitigating system PI?
- Is the yellow band of a PI twice as important as its white band?
- Is a yellow finding from an SDP of equal significance as a finding that a PI is in its yellow band?

We appreciate that judgments such as "of equal significance" and "twice as important" are subjective. Our argument is that attempting to answer questions such as these removes a good deal of the subjectivity and, in fact, will be very helpful when the thresholds are determined. This argument acquires additional significance in the present case in which the action matrix does not represent the judgments of a single individual but those of the agency. In other words, communication among the experts who make these judgments would be enhanced.