Performance Indicator Program Evaluation

Scope and Objectives—The staff of the U.S. Nuclear Regulatory Commission (NRC) evaluated the Reactor Oversight Process (ROP) performance indicator (PI) program in accordance with Inspection Manual Chapter 0307, "Reactor Oversight Process Self-Assessment Program." The staff used self-assessment metrics, external feedback through a *Federal Register* solicitation for comments, and periodic public meetings with industry and other stakeholders to evaluate the effectiveness of the PI program in fulfilling the regulatory principles of being objective, risk-informed, understandable, and predictable, as well as ensuring safety, openness, and effectiveness.

As a result of lessons learned, the staff has identified a number of issues and actions regarding the PI program. Enclosure 5 provides a summary of the status of these ongoing issues and actions, which are discussed in further detail below. The annual ROP performance metric report, available through the Agencywide Documents Access and Management System (ADAMS) provides the data and staff analysis for each of the program area metrics (reference ADAMS Accession No. ML060590135).

Summary of Previous Self-Assessment—In SECY-05-0070, "Reactor Oversight Process Self-Assessment for Calendar Year 2004," issued April 25, 2005, the staff discussed the history of the development of the PI program and provided an analysis of its historical results. The staff concluded that the PI program had not contributed to the early identification of poorly performing plants to the degree envisioned by the staff. The historical results showed that the percentage of greater-than-green PIs was 1.18 percent in calendar year (CY) 2000 and 0.47 percent in CY 2004, and that eight PIs had been all green following the first year of full ROP implementation. The staff discussed process issues that have delayed resolution of many frequently asked questions (FAQs), described the staff's actions to improve several PIs, and documented issues with several other PIs that need simplification or clarification. Overall, the staff stated that although the PI program continues to provide the U.S. Nuclear Regulatory Commission (NRC) with objective indicators regarding plant performance, and in some areas has focused licensee attention, thereby contributing to improved performance, the staff and some public stakeholders had become increasingly concerned with the lack of timeliness and inefficiency of the FAQ process, and with the capability of the current PIs to contribute to the identification of declining performance. Accordingly, the staff committed to engage senior industry management to define actions to address these issues.

PI Results—In CY 2005, the percentage of greater-than-green PIs continued to decline, to a value of 0.34 percent (although the number of PIs that crossed the green-white threshold increased from 8 to 11). The initiating events PIs continue to produce the most greater-than-green PIs (five), followed by four in the mitigating systems area and two in the emergency preparedness area. Nine out of the 15 nonsecurity PIs remained all green throughout CY 2005. Table 1 shows by PI and year the number of times any plant crossed from the green band into the white band.

When the PI program was developed, the green-white thresholds were set using industry performance data from 1995 to 1997 such that about 5 percent of the plants were expected to exceed the green-white threshold—that is, be designated either white, yellow, or red—for each

PI in the mitigating systems and initiating events cornerstones. Data from 1995 to 1997 were used to establish the thresholds because they were the most recent data available and the NRC considered industry performance in that period acceptable for the purposes of establishing ROP thresholds. The green-white thresholds for the barrier integrity PIs were set at 50 percent of the technical specification limit, and expert panels established the thresholds for the other PIs.

TABLE 1
NUMBER OF TIMES EACH PI
CROSSED THE THRESHOLD FROM GREEN TO WHITE

PI *	CY 2000 # Plants	CY 2001 # Plants	CY 2002 # Plants	CY 2003 # Plants	CY 2004 # Plants	CY 2005 # Plants	Total Plants
IE01	3	3	3	9**	4	3	25
IE02	3	0	0	1	1	1	6
IE03	3	1	1	0	0	1	6
MS01	5	1	1	1	1	0	9
MS02	6	1	1	0	0	2	10
MS03	4	2	2	1	0	2	11
MS04	0	0	0	0	0	0	0
MS05	3	0	0	0	0	0	3
BI01	0	0	0	0	0	0	0
BI02	2	2	0	1	1	0	6
EP01	1	0	1	0	0	2	4
EP02	0	0	0	0	0	0	0
EP03	6	0	0	0	1	0	7
OR01	1	0	0	2	0	0	3
PR01	0	0	0	0	0	0	0
	37	10	9	15	8	11	90

* IE01—Unplanned Scrams per 7000 Critical Hours; IE02—Unplanned Scrams with Loss of Normal Heat Removal; IE03—Unplanned Power Changes per 7000 Critical Hours; MS01—Emergency Alternating Current Power Systems Unavailability; MS02—High-Pressure Injection Systems Unavailability; MS03—Heat Removal Systems Unavailability; MS04—Residual Heat Removal Systems Unavailability; MS05—Safety System Functional Failures; BI01—Fuel Cladding Activity; BI02—Reactor Coolant System Integrity; EP01—Drill/Exercise Performance; EP02— Drill Participation; EP03—Alert and Notification System Reliability; OR01—Occupational Exposure Control Effectiveness; PR01—Radiological Effluent Occurrences

** The Northeast blackout in 2003 tripped nine nuclear units, only one of which crossed the white threshold due to that event.

In the first year of the ROP, each of the initiating events and mitigating systems PIs except MS04 (residual heat removal systems unavailability) identified from 3 to 6 percent of the plants

as outliers (those that crossed the green-white threshold), which was close to the expected value. However, in CY 2001, all of the initiating events and mitigating systems PIs except unplanned scrams declined significantly more than expected and have continued to decline. Improved industry performance explains some of this decline, but the staff remains concerned with the capability of the current PIs and PI thresholds to contribute to the identification of declining plant performance.

The PI program is a voluntary program and has no associated regulatory requirements. A standing working group with representatives from both the NRC and industry meets monthly to address issues associated with the program. In general, changes to the program need the agreement of both the industry and the NRC.

As committed to in last year's Commission paper, the staff engaged senior industry management to define actions to address these issues. Although the industry was open to future discussions on changes to the PI program, it made a number of points that the staff acknowledged, including the following:

- Since the staff deemed industry performance from 1995 to 1997 acceptable at the time of ROP development, any changes to PI thresholds could be seen as "ratcheting" industry performance beyond those agreed-upon levels.

- The staff should not rely on PIs as its main tool to detect declining performance; the inspection program considers a wider variety of areas, including operator actions, which provides the staff with better overall performance insights.

- The current PI definitions and thresholds have caused licensee actions that have improved overall safety; PI data and inspection program documentation do not capture many licensee actions.

- New PIs will need thorough review to ensure that they include no unintended consequences.

However, the staff continues to believe that the PI program should provide more input to the inspection program and help to identify declining plant performance. Therefore, the staff will continue to work with the industry to seek improvements to the program. The staff believes that this effort will require continued senior NRC and industry management involvement.

FAQ Process/Interpretation Issues—As discussed in last year's Commission paper, the staff committed to address issues related to the timeliness and efficiency of the FAQ process. FAQs are specific questions raised about the interpretation of PI implementation guidance and often influence the color of a PI at the plant submitting the FAQ. The staff previously noted that many PIs lack clear, concise guidance, which contributed to the timeliness and efficiency issues. The staff also had trouble handling a number of potentially white PIs with interpretation issues. The staff stated that these issues are often caused by different interpretations of the PI guidance document, Nuclear Energy Institute (NEI) 99-02, "Regulatory Assessment Performance Indicator Guidelines," and the joint NRC/industry working group had difficulty agreeing on which events or conditions the PI calculation should include. To address issues for

which it is clear that the staff and industry will not reach agreement, the working group established a process to come to a final decision—the issue will be raised to a division director in the Office of Nuclear Reactor Regulation, whose decision will be final.

After implementing the new process in CY 2005, the backlog of FAQs decreased dramatically from 24 open FAQs in April 2005, to 2 open FAQs as of January 1, 2006. This change, and the staff's increased focus on timeliness, has improved overall FAQ efficiency and effectiveness. However, a number of issues remain that impact the effectiveness of the PI program, including the following:

- As a result of the number of revisions to the guidance document to incorporate FAQs, the document now may not be as clear and concise as it should be.

- The staff and the industry disagree about crediting operator actions. The industry wants to credit operator actions for the event as it occurred. The staff believes that, except for a few very simple actions (such as pushing the manual start button when an automatic start fails), operator actions cannot be credited. The objective of the PIs is to measure equipment performance, not human performance.

- NEI 99-02 interpretation issues continue to remain a staff concern. Since the staff conducts PI verification inspections once a year, a potential disagreement might not surface for up to a year or more after the event in question, so that by the time the question reaches the working group as an FAQ, it is already untimely.

- Licensees have performed lengthy engineering evaluations to demonstrate that an event or condition did not render a system unavailable, and at times these have not been timely in accordance with the guidance document.

To address some of the concerns described above, the staff plans to work with the industry to review each PI definition and supporting information in the NEI guidance document. This effort is not intended to change definitions, but rather to make the document more readable, more concise, and not subject to as much interpretation. This effort, however, will not totally address all of the staff concerns mentioned above regarding efficiency and effectiveness. To address these, the staff will continue to work with the industry and advocate for change in PI definitions and guidance as discussed previously.

If the staff believes that a licensee has not appropriately reported a PI, the staff has the option to conduct an inspection of the PI using Inspection Procedure 71150, "Discrepant or Unreported Performance Indicator Data." Under this process, the staff may declare a licensee's PI data report invalid and color the associated PI gray until the staff determines the correct color through inspection. In 2005, the staff implemented the discrepant PI process twice. The Davis-Besse alert and notification system PI data were determined to be discrepant in January 2005. In November 2005, the staff identified the Waterford 3 high-pressure safety injection system PI as potentially discrepant; however, the NRC has not completed the discrepant PI inspection, which is scheduled during the first quarter of 2006. This process takes time to schedule and implement and is relatively costly in resources. It illustrates the value of PI definitions that are clear and predictable and that require little evaluation or analysis to correctly report.

PI Improvements

The following improvements to the PI program were in process during CY 2005 and are listed in the order of priority.

(1) Mitigating Systems Performance Index (MSPI)—The staff and industry developed the MSPI to address known problems with the safety system unavailability PIs. MSPI is a complex, risk-based process that combines component reliability and availability with plant-specific probabilistic risk assessment (PRA) information to arrive at a single performance index for the monitored system. Specific individual plant design influences MSPI significantly. Since conclusion of the MSPI pilot in 2003, the staff and industry have continued to work to finalize the technical guidance needed for implementing MSPI, define and address a minimum level of PRA quality needed for MSPI, develop the databases and software necessary for each licensee to be able to implement MSPI, and resolve issues identified throughout the development and review processes. The staff plans to replace the safety system unavailability PIs with MSPI at the beginning of April 2006, with the first submittal (including three years of data) in July 2006.

NEI has the lead responsibility for working with all operating reactor licensees to prepare them to implement MSPI. NEI has sought to accomplish this goal by hosting several technical workshops over the course of the past several years. The final MSPI workshop was held in early February 2005. One of the principal topics discussed in the later MSPI workshops was PRA quality. In December 2004, a staff/industry MSPI task group recommended a minimum level of PRA quality that each licensee PRA must meet for MSPI purposes. NEI, licensee representatives, and the NRC staff accepted the recommendations of the task group during a public meeting in January 2005. In the March 2005 ROP working group meeting the industry indicated that it could not meet the PRA quality criteria previously agreed upon by the original MSPI implementation date (January 2006). NEI subsequently proposed a change to the PRA quality activities to be completed before MSPI implementation. In the alternate approach, industry would conduct an MSPI component comparison study using predetermined criteria that would identify monitored component outliers in terms of risk significance (primarily pumps and emergency diesel generators). Licensees who were identified as having outliers would need to address why the associated component is an acceptable outlier, or take action to resolve the issue of concern.

Over the past 12 months, the Office of Research, with support from the regions, conducted an independent component risk comparison study using industry data to verify and validate the industry comparison study. The study identified approximately 265 candidate component outliers whose contribution to plant risk was significantly different than the risk value as calculated using the staff's updated standardized plant analysis risk (SPAR) models. The staff has managed to resolve most of these outliers by pinpointing the specific differences and limitations of the SPAR models when compared to the plant PRAs, as well as highlighting a number of accuracy and modeling issues with plant PRAs. In parallel with this effort, the staff reviewed each licensee's MSPI basis document during the fall of 2005. Based on this review, the staff and the industry concluded that a significant number of licensees were not ready to implement MSPI by January 2006. Therefore, the ROP working group decided to delay MSPI implementation until April 2006.

The staff had previously completed its review of the Institute of Nuclear Power Operations (INPO) Consolidated Data Entry (CDE) Program and found that it is consistent with the staff's

needs for licensee data used in various NRC programs. The CDE Program was used successfully in CY 2004 to gather and submit PI data to the NRC. The staff will complete its review of the CDE Program for the MSPI indicators and will test the submittal and posting process prior to MSPI implementation.

(2) Safety System Functional Failures (SSFFs)—The benchmarking analysis of the proposed PIs in 1999 showed that this PI identified all the watch list plants, all the declining trend plants, and some of the below average plants. This PI counts all events or conditions that could have prevented the fulfillment of the safety function of structures or systems needed to shut down the reactor, remove residual heat, control the release of radioactive material, or mitigate the consequences of an accident. Title 10, Section 50.73(a)(2)(v), of the *Code of Federal Regulations* (10 CFR 50.73(a)(2)(v)) requires such events to be reported in licensee event reports (LERs). Copies of all LERs received by the NRC are sent to the Idaho National Laboratory (INL), which is under contract to the NRC to read and code all LERs for use in the Industry Trends Program, which includes SSFFs. Every quarter licensees report to the ROP PI program the number of events or conditions that the licensees determine could have prevented the fulfillment of any of the above safety functions. However, the number of events reported by licensees is less than the number of events captured as SSFFs by INL.

To examine this difference, several industry members of the ROP working group volunteered to review licensee event reporting from 1999 to 2003. They determined that there is some confusion regarding the reporting requirements under 10 CFR 50.73 that needs to be addressed. The group recommended a number of possible changes to NUREG-1022, "Event Reporting Guidelines 10 CFR 50.72 and 50.73," and the process INL uses to categorize LERs.

The staff generally agrees that there can be differences of opinion on the reporting requirements in NUREG-1022, which states that the standard for reporting is "a reasonable expectation of not fulfilling the safety function." The staff and industry have differing views on what is "reasonable." Secondly, a number of events have occurred that are not explicitly covered in the guidance document, leaving room for interpretation. The staff believes that the NEI 99-02 guidance is clear that the PI includes a wide variety of events or conditions, ranging from actual failures on demand to potential failures attributable to various causes, including environmental qualification, seismic qualification, human error, design or installation errors, etc. Many SSFFs do not involve actual failures of equipment. Because the contribution to risk of the structures and systems included in the SSFF varies considerably, and because potential as well as actual failures are included, it is not possible to assign a risk significance to this indicator.

This differs somewhat from the NUREG guidance, but it is these potential failures that were reported more frequently prior to ROP implementation but have not been reported as consistently since ROP implementation. To address the issues with this PI, the staff plans to review the data in this area further and continue to discuss this PI with the industry. The staff also plans to evaluate the NUREG for possible changes.

(3) Scrams with Loss of Normal Heat Removal—The backlog of FAQs for this PI was cleared in CY 2005. The staff used, for the first time, the dispute resolution process that had the NRC Director of the Division of Inspection Program Management (now the Division of Inspection and Regional Support) make the final determination. The staff used this process for six events. In addition, the NRC/industry task group reached agreement on the definition for a proposed replacement PI for monitoring complicated scrams. Data collection to establish thresholds is

underway. The staff believes that this new PI has the potential to be a leading indicator of declining performance in that a plant that has a history of complicated scrams may be more likely to have a risk-significant scram.

(4) Unplanned Scrams per 7000 Critical Hours—In CY 2005, the staff proposed to change this PI to unplanned scrams per 7950 hours. During PI development in the late 1990s, the PI was defined as unplanned scrams per 7000 critical hours because the industry average plant availability in 1995 through 1997 was 80 percent, or 7000 critical hours in four quarters. The green-white threshold is met if a plant, in any consecutive four quarters, has either four or more scrams, or three scrams and fewer than the industry average number of critical hours, rounded off to the nearest one-tenth of one percent. Today the industry average availability is just under 91 percent, which is equivalent to about 7950 hours per year of operation. To maintain the basis for identifying outliers with four or more scrams in four quarters (and those with three scrams and fewer than the industry average availability), the staff is considering whether the numerator should reflect the industry's current average availability in one year. The industry believes that this change represents an increase in expected plant performance, above that which was deemed adequate in the 1995–1997 timeframe. Therefore, the industry is opposed to the change. The staff also proposed to change the red threshold for this PI to a deterministic value, down from the current risk-based threshold of 25, as recommended by the Advisory Committee on Reactor Safeguards. The staff will discuss these changes further with stakeholders during the monthly ROP meetings.

(5) Unplanned Power Changes per 7000 Critical Hours—In CY 2005, the staff proposed to change this PI to unplanned power changes per 7950 critical hours for the same reason as described for the unplanned scrams PI. The staff is also considering changing the definition of "unplanned" from at least 72 hours in advance to at least 4 weeks in advance to align with the similar indicator used by the World Association of Nuclear Operators (WANO) and to avoid problems with the 72-hour requirement that have resulted in many FAQs.

In addition, in CY 2005 the staff proposed to include in the definition for this PI NRC notices of enforcement discretion (NOEDs) that preclude power changes. This potential change was not discussed in much detail, and, because of the small number of NOEDs currently issued per year, the impact on the PI would be small. However, since both internal stakeholders and members of the public proposed this change, the staff will continue to keep this issue open and engage industry further.

(6) Reactor Coolant System Leakage—The Davis-Besse Lessons Learned Task Force recommended that the staff continue ongoing efforts to review and improve the usefulness of the barrier integrity PIs and evaluate the feasibility of establishing a PI that tracks the number, duration, and rate of primary system leaks that have been identified but not corrected. The NRC/industry task group agreed that this approach was likely not feasible but did, however, agree that this PI should monitor unidentified leakage (instead of identified leakage) averaged over an appropriate time interval to identify baseline values and trends. The Westinghouse and Boiling-Water Reactors Owners Groups have also established a working group to examine leakage monitoring. The owners groups are currently collecting data for a revised PI and expect to have results in the fall of 2006. If these efforts identify a feasible method for improving the PI, it will be presented to the ROP working group for further discussion.

(7) Reactor Coolant System Activity—The staff continues to pursue the use of the WANO fuel reliability indicator. INPO staff members have stated that they have encountered problems with this indicator that are similar to those the NRC staff has had with its PI. Therefore, INPO does not recommend that the NRC adopt the WANO indicator. This issue is currently on hold while the staff addresses higher priority items.

(8) Containment Leakage—The staff removed this PI from the ROP PIs after the pilot program for the following reasons: (1) the test methods used and data collected across the industry lack uniformity because licensees may choose between two methods for performing leak rate tests, only one of which requires recording the as-found leakage, (2) the tests are normally conducted during refueling outages, so the data are end-of-cycle numbers which may or may not be indicative of the worst-case leakage in the previous operating cycle, and (3) licensees are required to restore containment leak rates to within acceptable limits prior to restart.

Nevertheless, this PI may provide some value if it encourages licensees to adopt a more uniform test methodology, and even a backward look at containment integrity could be of value in identifying recurrent issues. For these reasons, the staff plans to discuss the value of this PI with external stakeholders.

(9) Potential New Indicators—The staff will also consider the feasibility of new indicators. One readily available candidate, used in the NRC's Industry Trends Program, is safety system actuations. Safety system actuations are under consideration as a new indicator because the Industry Trends Program has found that their number is increasing and exceeded the short-term prediction limit in fiscal year 2005.

The staff also plans to look for indicators for cross-cutting issues, as recommended by the Advisory Committee on Reactor Safeguards and other stakeholders, but this effort is on hold while the staff addresses higher priority issues.

PI Program Performance Metrics—The four metrics that met their established criteria are (1) effectively responding to questions regarding interpretation of PI guidance, (2) timely indication of declining safety performance, (3) timely PI data reporting and dissemination, and (4) clarity of PI guidance. The three metrics that did not meet their established criteria are (1) achieving consistent results given the same guidance, (2) obtaining insights from the PI program to help ensure plant safety, and (3) stakeholders perceiving appropriate overlap of PIs and the inspection program.

The staff tracks significant discrepancies and discrepant PIs reported in each quarter to determine whether consistent results are achieved given the same guidance. In the second and fourth quarters of CY 2005, the number of significant discrepancies or discrepant PIs increased notably in comparison to previous years; five discrepancies occurred in CY 2005 as compared to two discrepancies in both CY 2001 and CY 2002 and none in CY 2003 and CY 2004. The other two missed PI metrics resulted from the negative public perception noted in the external survey responses as described below. As previously discussed, the staff continues to evaluate several PIs, with inputs from internal and external stakeholders, in an effort to improve their effectiveness at identifying poor performance.

Stakeholder Survey Results—The staff did not conduct an internal survey in CY 2005; therefore, the input to this discussion came solely from the external survey conducted in

October 2005. Participants in the external ROP survey included nine industry representatives, four State or local government agencies, seven private citizens or public interest groups, and one anonymous stakeholder. The external survey revealed a sharp difference in opinion between public interest groups/members of the public and industry groups regarding the effectiveness of the PI program. The public generally believed that the PIs have become less useful and provide little information on plant performance because they are almost always all green, while industry comments concluded that the PIs provide useful indicators of plant safety. Industry groups generally felt that the degree of overlap between the PI program and the inspection program was appropriate, or perhaps somewhat excessive, while public satisfaction in the degree of overlap decreased, largely because the PIs are almost always all green. Enclosure 6 provides more detail on the results of the external survey.

Self-Assessment Conclusions—Based on the metric results, stakeholder feedback, and other lessons learned and evaluations, the staff concludes that the PI program has successfully met the goals and intended outcomes of the ROP. The PI program has generally fulfilled the regulatory principles of being objective, risk-informed, understandable, and predictable, and it has accomplished the three applicable NRC strategic goals (ensuring safety, openness, and effectiveness). The timeliness and efficiency of the FAQ process appears to have improved significantly. Although the PI program continues to provide the NRC with objective indicators regarding plant performance, and in some areas has focused licensee attention and contributed to improved performance, the staff and some public stakeholders remain concerned with the capability of the current PIs to contribute to the identification of declining performance. As a result, the staff are in the process of revising several PIs. The staff plans to continue to work with the industry to revise and/or introduce other PIs to improve the program's effectiveness in contributing to the identification of declining performance. The staff believes that this effort will require continued senior NRC and industry management involvement.