

ROP Performance Metrics

The Reactor Oversight Process (ROP) performance metrics utilize objective measures and pre-determined criteria to monitor the performance of the ROP as described in Inspection Manual Chapter (IMC) 0307, "Reactor Oversight Process Self-Assessment Program." These metrics rely on information from various sources, including the reactor program system (RPS), the inspection program, periodic independent audits, stakeholder surveys, and public comment. Metrics have been developed to monitor each major component of the ROP, as well as metrics of a more general nature intended to gauge overall ROP performance.

Data is collected on a quarterly basis, as applicable, and is compared to pre-established criteria for analysis. In most cases, success is defined as a steady or improving trend. Quantitative success criteria for many of the performance metrics has not been developed due to the infancy of the ROP and the lack of data needed to establish objective criteria. For these metrics, baseline data was collected and continues to be used to monitor trends and develop criteria for the future, as appropriate.

The Nuclear Regulatory Commission (NRC) solicited comments on the third year of ROP implementation from external stakeholders in a *Federal Register* notice published on November 22, 2002. Participants included a member of the general public, four public interest groups, nine utilities and utility interest groups, four State regulatory agencies, and one law firm. Additionally, the NRC conducted an internal survey to obtain feedback on the ROP via a web-based questionnaire in November 2002. The detailed analysis of the internal survey is included in Attachment 5 to this paper. The detailed comments from the external survey are consolidated into a summary document (ADAMS accession number ML030620007) and a discussion of this survey is included in Attachment 4 to this paper. Staff analysis of the specific responses is also included in the applicable portions of the program area discussions in this paper as well as in the ROP performance metric report in Attachment 3.

The majority of metrics met their established criteria. All metrics in the inspection and assessment areas met their criteria, but some metrics in the Performance Indicator (PI), Significance Determination Process (SDP), and the overall ROP areas did not meet their success criteria. The staff's corrective actions to address these issues are discussed below and in the applicable program area discussions in Attachment 1.

One of the seven PI metrics was determined to not meet its established criteria based on the negative perception by the public that the PI program may adversely impact plant safety (PI-4). Several survey respondents mentioned concerns that the Davis-Besse reactor head issue was not identified by the ROP and that there was too much focus on risk significance as support for their negative statements. To address these concerns, the staff plans to enhance the barrier integrity PIs to better detect unidentified leakage as recommended by the Davis-Besse Lessons Learned Task Force (DBLLTF), and is currently evaluating the need and feasibility for a public workshop in calendar year (CY) 2003 to address several of the common concerns noted by both the internal and external stakeholders.

Of the nine metrics counted for the SDP, four did not meet their established criteria. Two of these unsuccessful metrics resulted from the negative perception regarding the inspection staff's proficiency in using the SDP (SDP-3) and that the SDP results translate to the same level of significance for all cornerstones (SDP-5). The other two were based on SDP timeliness (SDP-8) and the accuracy of results communicated to the public (SDP-9). The timeliness issue continues to be pursued by the staff and improvements are expected as the process is refined.

The accuracy issues resulted from either untimely posting of information to the Web or inaccurate labeling of a finding allowing it to be double-counted. These issues, along with others, continue to be addressed via the SDP Improvement Initiative as discussed in further detail in the SDP program area discussion in Attachment 1.

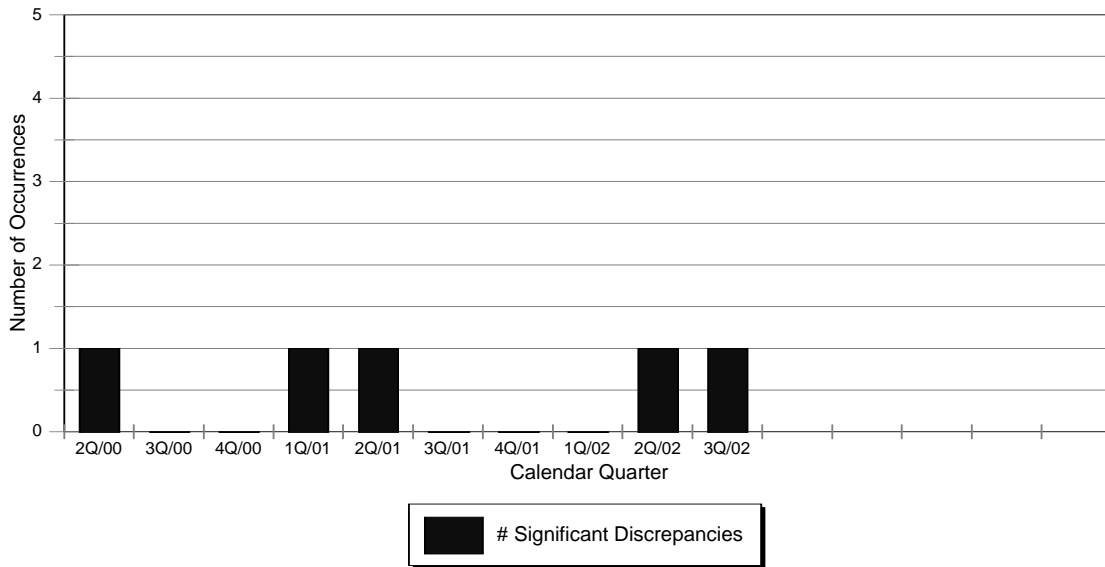
Of the nineteen overall metrics established for the ROP, four were determined to not meet the established criteria. Based on the recommendations and programmatic deficiencies noted by the DBLLTF, the staff concluded that the metric established to measure whether there are any programmatic voids in the ROP (O-9) was not met. In addition, three metrics gauging the public's perception of the ROP were determined to not meet the established criteria. These metrics include whether the ROP maintains safety (O-7), whether the ROP is effective, efficient, and realistic (O-11), and whether the ROP results in unintended consequences (O-19). The ROP's ability to maintain safety was questioned by some stakeholders because the ROP did not successfully identify the vessel head degradation at Davis-Besse. In addition, a common concern among most respondents regarding the effectiveness of the ROP was the efficiency and realism of the SDP. The staff plans to implement the recommendations of the DBLLTF and the SDP Task Group, and is also evaluating the need and feasibility for a public workshop in CY 2003, to address several of the common concerns noted by both the internal and external stakeholders.

The analysis of the metrics provided insights into other program areas in need of improvement. The detailed metrics and their analysis are provided on the following pages.

PI-1 Consistent Results Given Same Guidance

Definition: Independently verify PIs using Inspection Procedure (IP) 71151, “PI Verification.” Count all PIs that cross a threshold because of discrepancies as noted in the resultant inspection report. Licensees are requested per Nuclear Energy Institute (NEI) 99-02 to report changes to PI colors as soon as practical upon discovery via a “mid-quarter” report and to annotate in the comments field an explanation for the change.

Criteria: Use the first year of data as a benchmark for future comparison and to establish



acceptable range of variability.

Comments: The graph represents the number of significant discrepancies reported for each quarter. Significant discrepancies are issues identified by the NRC during a PI verification inspection that caused the PI to cross a threshold.

Analysis: Two significant discrepancies (2nd and 3rd quarter of 2002) were identified through PI verification (IP 71151, “Identification and Resolution of Problems”) inspections conducted during the assessment period. The discrepancies occurred at two different plants.

At one plant, the inspectors identified that the licensee incorrectly assessed the accuracy of some emergency preparedness initial notification forms. Once the licensee reassessed these forms and resubmitted their PI data, the Drill and Exercise Performance PI changed from green to white. At another facility, inspectors questioned the licensee’s practice of not counting Emergency AC Power (EAC) unavailable hours during the performance of monthly Emergency Diesel Generator (EDG) surveillance tests. The issue was discussed by the ROP Working Group in a public meeting and it was decided that the EAC was unavailable during the monthly

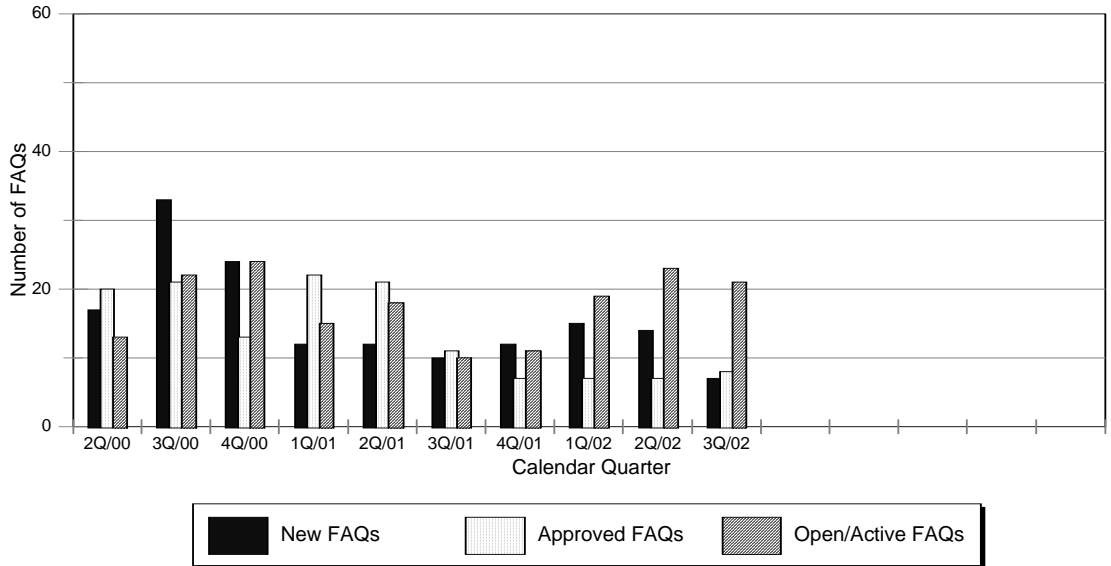
tests. The licensee thereupon entered the appropriate hours into the EAC Safety System Unavailability (SSU) performance indicator calculation and resubmitted their PI data. Although the discrepant unavailability time that was added back was not a significant contributor to the overall EAC unavailability, it resulted in the performance indicator crossing the white threshold starting in the fourth quarter of 2001.

The number of these discrepancies remains very low. When all data, starting with the full implementation of the ROP, is taken into account, this reflects a stable trend.

PI-2 Questions Regarding Interpretation of PI Guidance

Definition: Quarterly, count the number of frequently asked questions (FAQs).

Criteria: Expect low numbers (but not as low as metric PI-1), with a stable or decreasing trend.



Comments: Each quarter represents the total number of new FAQs introduced and approved during the ROP NRC/Industry Working Group meetings held during the respective quarter. This metric was revamped after insights were gained from the first year of full implementation of the ROP. The improvements made to this metric provide for a more timely and accurate account of FAQs. Since this metric was reconstructed from historical data, the second and third quarter of 2000 contain estimates (some FAQ logs were unavailable).

Analysis: Interpretation questions regarding the PI guidance in NEI 99-02 took an upward trend during the initial stages of the ROP. This upward trend was anticipated; however, as NRC inspectors and licensees became more familiar with the guidance, and as additional guidance was provided to clarify NEI 99-02, a lower and generally stable number of questions required evaluation. Recently (1st and 2nd quarter 2002), the total number of open/active FAQs has trended slightly up. This trend is due in part to the resources diverted to the pilot testing of the mitigating system performance index (MSPI) and to the ongoing challenges with the Safety System Unavailability, Scrams with Loss of Normal Heat Removal, and Unplanned Power Changes PIs, which are described below. Preliminary data for 4th quarter 2002 indicates that the total number of open/active FAQs appears to be decreasing.

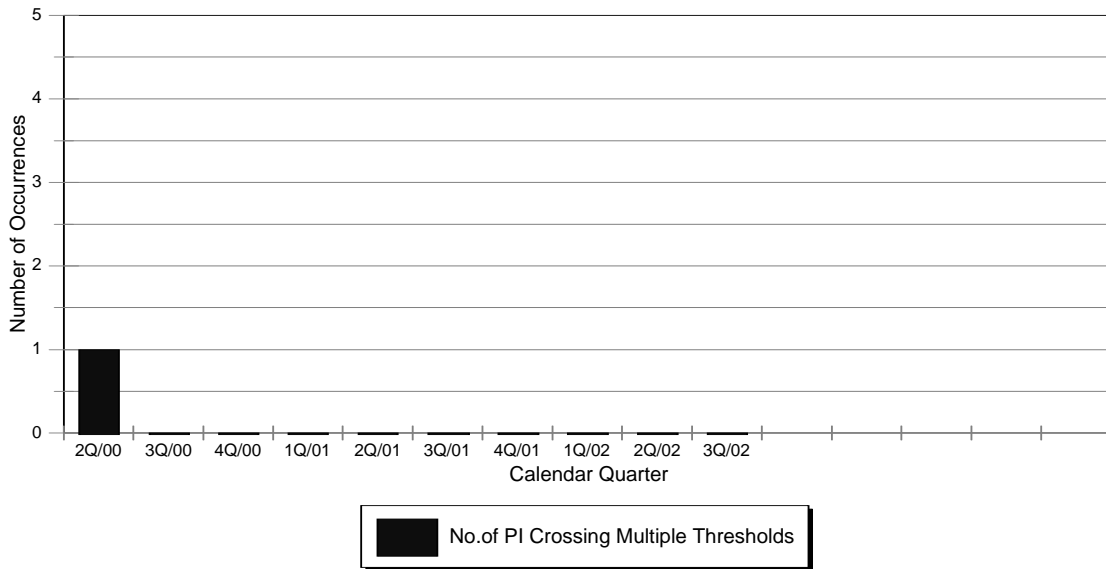
The largest number of FAQs continue to reflect the challenges with the mitigating systems cornerstone (unavailability). Pilot testing of the proposed MSPI developed to replace the current Safety System Unavailability PIs is currently in progress. A significant number of FAQs

were related to the initiating event PIs of Scrams with Loss of Normal Heat Removal and Unplanned Power Changes. These FAQs have resulted in extended discussions during the ROP Working Group public meetings. The staff is analyzing options to clarify these PIs and has asked stakeholders for input.

PI-3 Timely Indication of Declining Safety Performance

Definition: Quarterly, track PIs that cross multiple thresholds (e.g., green to yellow or red). Evaluate and characterize these results to allow timely indication of declining performance.

Criteria: Expect low numbers (near zero).



Analysis: There were no occurrences of PIs crossing multiple thresholds during this assessment period. For the given parameters that have been included in the PIs, the PIs appear to provide timely indication of declining performance.

PI-4 Minimize Potential for Licensee Actions Taken in Response to the Performance Indicator Program That Adversely Impact Plant Safety

Definition: Survey stakeholders regarding PIs driving undesirable decisions. This question will be included in the overall *Federal Register* notice.

Criteria: Expect low numbers of unintended consequences reported, with a stable or decreasing trend.

Analysis: All of the utility/utility group respondents stated, or endorsed NEI's comment, that the PI program together with the inspection program provides incentives to minimize the potential for licensees to take actions that adversely impact plant safety.

However, all of the public interest group respondents stated that the PIs do not minimize the potential for licensees to take actions that adversely impact plant safety. Their concerns included:

- that the PI program did not identify, in the case of Davis-Besse, significant programmatic breakdowns in the safety evaluation and corrective action processes
- the industry campaigns to change PI thresholds, definitions, etc., to ensure PIs remain green
- green PIs seem to be giving plant owners and the NRC a false sense of safety
- conditions or events are dealt with in an atomistic fashion rather than in a holistic way (e.g., every indication of the Davis-Besse reactor pressure vessel head issue was of itself of very low safety significance, none of which triggered enhanced NRC oversight until the situation became extreme)
- when coupled with other "risk-informed" initiatives, has allowed licensees to delude themselves into ignoring safety problems

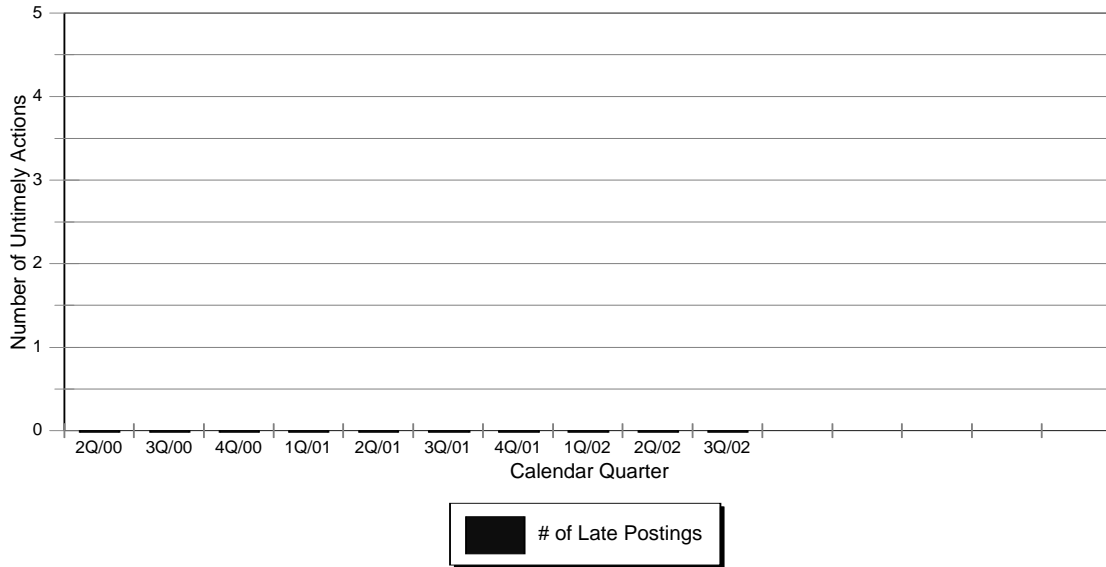
Two of the State regulatory agencies stated that there was still a small potential for licensees to inadvertently take actions that might adversely impact plant safety. The other participants did not directly respond to this item.

The criteria for this metric has not been met, primarily due to the responses received from the public interest groups.

PI-5 Timely PI Data Reporting

Definition: Within 5 weeks of the end of each calendar quarter, track (count) late PI postings on the NRC’s external Web site.

Criteria: Expect a low number (near zero) of late PI postings on the NRC’s external Web site.



Analysis: There have been no late PI data submittals from licensees, or subsequent postings to the web page, since the inception of the ROP.

PI-6 Stakeholders Perceive Appropriate Overlap of Inspection Program and PIs

Definition: Survey stakeholders' perceptions of overlap between PIs and the Inspection Program. This question will be included in the survey for internal stakeholders and the *Federal Register* notice for external stakeholders.

Criteria: Expect a low number of negative comments, with a declining or stable trend in the number of negative comments received.

Analysis:

Internal Survey

Listed below are the staff's responses to the following statement: "The Performance Indicators provide an appropriate level of overlap with inspection program."

Strongly agree: 3.4 %
Agree: 60.3 %
Disagree: 17.7 %
Strongly disagree: . . 5.2 %
Unable to answer: . 13.4 %

Not including the "unable to answer" responses, 73.7% of the respondents agreed that the PIs provide an appropriate level of overlap with the inspection program. This result is similar to that received from the previous internal survey conducted in March of 2001, in which 74% of the respondents agreed.

External Survey

All of the utility/utility group responses to the external survey stated, or endorsed NEI's comment, that in general appropriate overlap exists between the PI program and the inspection program and further commented that if anything, there was excessive overlap (i.e., in the radiation protection and emergency preparedness areas already covered by PIs).

Responses to the external survey received from public interest groups all stated that there was not appropriate overlap between the PIs and the Inspection Program. Several of these respondents stated that there was too much of a focus on risk significance and one stated that there were not enough inspections. Several of the respondents mentioned the Davis-Besse reactor head issue as support for their statements. Two of the State regulatory agencies stated that there was an appropriate level of overlap, and one indicated that this item is not easily measured but it didn't identify issues related to the Davis-Besse reactor head. The other participants did not directly respond to this item.

The criteria for this metric has been met based on low number of negative comments and a stable perception regarding appropriate overlap.

PI-7 Reporting Conflict Reduction

Definition: Survey licensees and other external stakeholders regarding the perceived overlap between reporting requirements, such as those promulgated by the Institute of Nuclear Power Operations (INPO), the World Association of Nuclear Operators (WANO), and the Maintenance Rule. This question will be included in the Federal Register notice.

Criteria: Expect a low number of negative comments, with a declining or stable trend in the number of negative comments received.

Analysis: In last year's ROP annual assessment (for calendar year 2001), every respondent that specifically commented on this item indicated that in some manner conflicts exist, especially in the area of safety system unavailability within the Mitigating System Cornerstone.

In response to this year's external survey, every utility/utility group respondent commented, or endorsed NEI's comment, that there are differences in reporting and definitions among the ROP, WANO/INPO, and the maintenance rule. Many of the respondents noted that these differences are being addressed by the proposed Mitigating System Performance Index which is currently being pilot tested. These respondents further noted that industry is also working to reduce the unnecessary duplicative reporting with the introduction of the Consolidated Data Entry system being developed by INPO.

Several non-utility stakeholders responded that they could not comment on any items related to INPO and WANO since the INPO and WANO documents are not publicly available, but with respect to the maintenance rule, there was no undue conflict or unnecessary overlap. One non-utility stakeholder commented that the NRC should not care if WANO and INPO requirements are duplicative or not.

Although the utility respondents commented that differences exist between ROP, WANO/INPO, and the maintenance rule, the NRC and industry currently have improvements in progress to address these differences. This metric has been met based on the declining number of negative comments.

PI-8 Clarity of PI Guidance - NEI-99-02

Definition: Survey external stakeholders' perceptions regarding the clarity of the guidance contained in NEI 99-02. This question will be included in the Federal Register notice.

Criteria: Expect a low number of negative comments or examples of interpretation issues, with a stable or declining trend in the number of negative comments received.

Analysis: The vast majority of utility/utility group respondents commented, or endorsed the NEI comment, that the guidance contained in NEI 99-02 was generally clear and adequate. However, several respondents commented on issues relating to problems with definitions associated with the Scrams with Loss of Normal Heat Removal indicator and on timeliness of the Frequently Asked Questions process used to resolve interpretation issues.

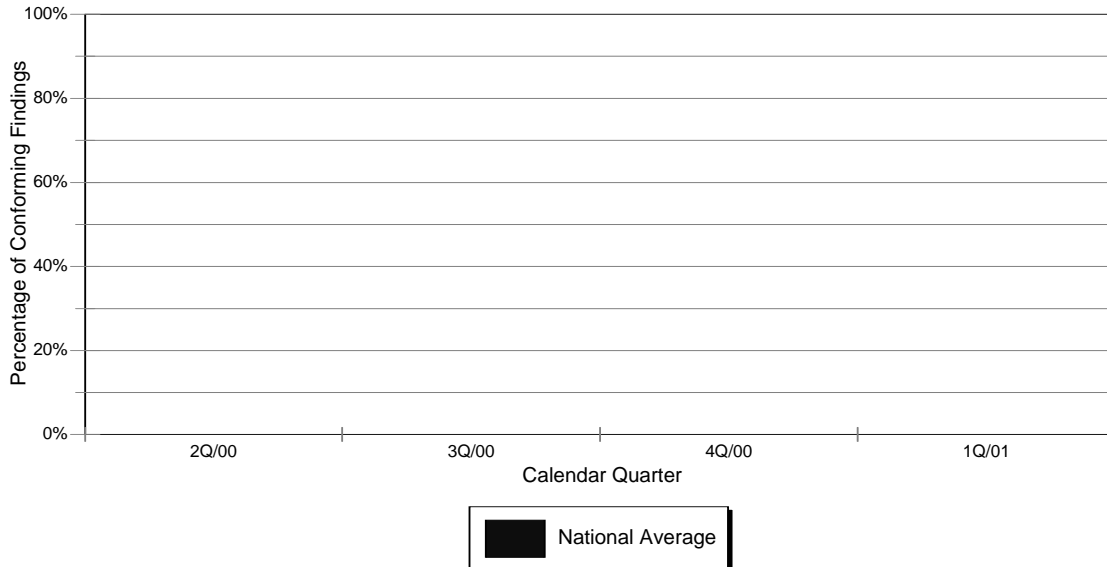
A few non-utility stakeholders commented directly on the clarity of the NEI 99-02 guidance. Two respondents considered in general that the guidance was helpful or clear (with one noting that the current PIs might not be appropriate or telling). One negative comment was received that indicated that there was too much room for interpretation in the guidance.

The criteria for this metric was met since a low number of negative comments or examples of interpretation issues were received.

IP-1 Percentage of Inspection Findings In Accordance With Requirements

Definition: Audit inspection reports in relation to program requirements (IMC 0612, “Power Reactor Inspection Reports”) for documenting green findings, greater-than-green findings, and violations. Report the percentage of findings that meet the program requirements. Each year, audit one resident/integrated report from each plant, 25 percent of all other baseline reports, and all reports resulting from inspections beyond the baseline program.

Criteria: Expect an improving trend in the percentage of findings documented in accordance with program requirements.



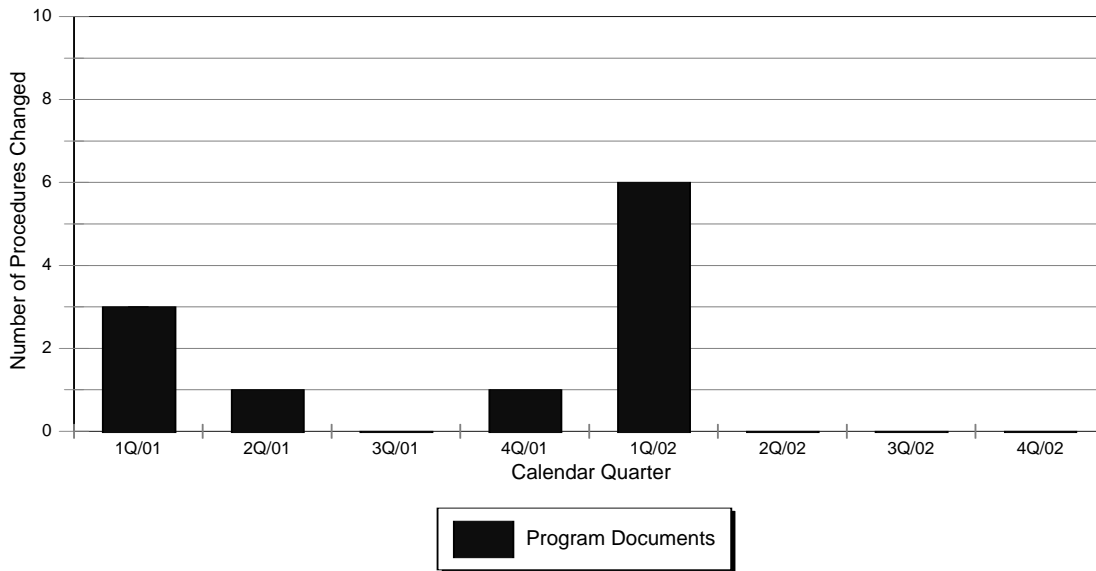
Comments: The Office of Nuclear Reactor Regulation (NRR) staff issued IMC 0612 on April 29, 2002, to improve the program guidance on documentation of inspection findings. After a brief training period, all regions implemented the new requirements of IMC 0612 in July of 2002. To allow inspectors and regional management to become more familiar with and implement the new requirements of IMC 0612 and time to issue a sample inspection report for regional use, the staff agreed to allow the first set of inspection report audits under IMC 0612 be conducted by regional personnel. The NRR staff will commence auditing the inspection reports in CY 2003.

Analysis: No data was available in CY 2002 due to program transition.

IP-2 Number of Baseline Inspection Procedures Significantly Changed

Definition: Review all issued changes to baseline inspection procedures and count those procedures whose scope or frequency of inspection changed, and count new inspectable areas that relate to risk-informing the inspection.

Criteria: Expect relatively few significant changes, with a stable or declining trend.

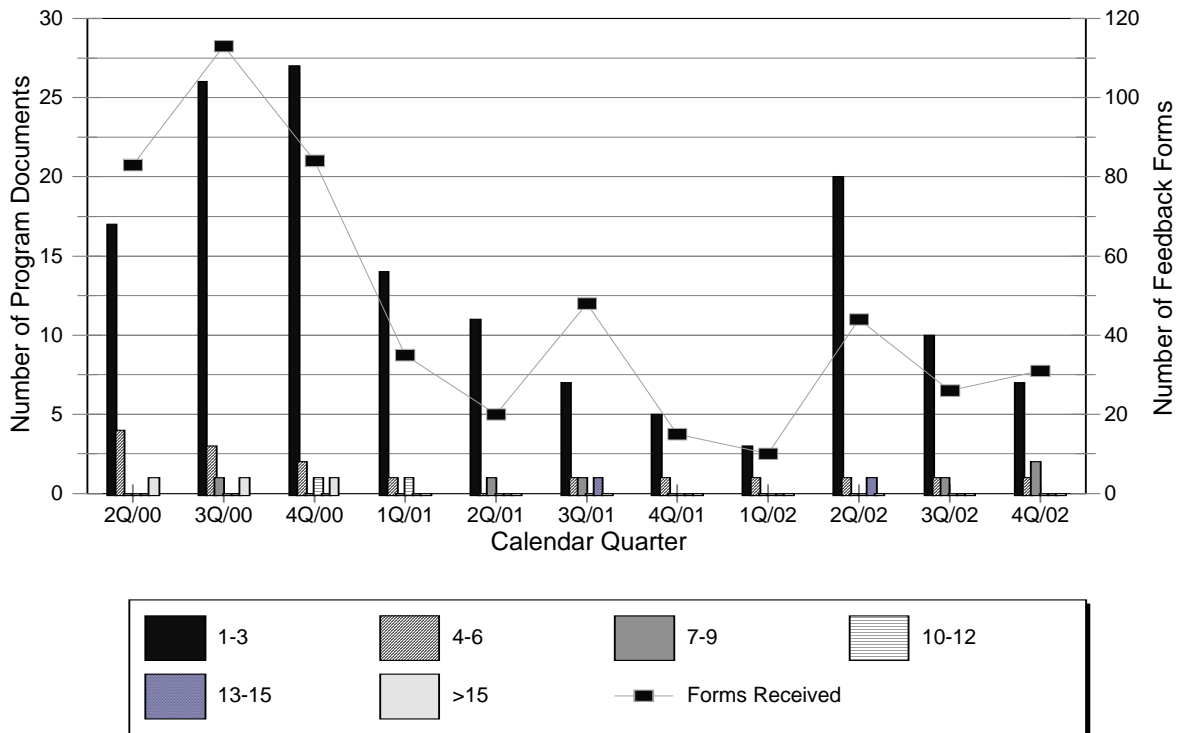


Analysis: There was a total of six changes in the first quarter of 2002 that affected either the scope or frequency of baseline inspection procedures. Inspection frequency of IP 71111.02, “Evaluation of Changes, Test, or Experiments,” and IP 71152, “Identification and Resolution of Problems,” was changed to biennial frequency and a security section was added to supplemental procedure IP 95003, “Supplemental Inspection For Repetitive Cornerstone, Multiple Degraded Cornerstone, Multiple Yellow Inputs, or One Red Input.” Additionally, a new inspection requirement was added to the radiation procedures IP 71121.03, “Radiation Monitoring Instrumentation,” IP 71122.02, “Radioactive Material Processing and Transportation,” and IP 71122.03, “Radiological Environmental Monitoring Program (REMP).” Although there was a sharp increase in the number of changes to the baseline inspection procedures in the first quarter of 2002, the aggregate number of significant changes during the years 2001 and 2002 remain nearly constant at around five. This metric was met based on the relatively stable trend.

IP-3 Number of Feedback Forms per Document

Definition: Count the number of feedback forms received for each program document each quarter. Use a histogram to chart the number of documents for which feedback forms were received. Highlight those documents against which the most forms are written.

Criteria: Expect a decreasing trend in the number of feedback forms received for program documents.



Analysis: The staff received 112 feedback forms during the 2002 calendar year. Approximately 60% of all feedback forms received during 2002 were related to issues in the following areas: (1) Operating Reactor Assessment Program (IMC 0305, “Operating Reactor Assessment Program”); (2) Performance Indicator Program (IMC 0608, “Performance Indicator Program”); (3) Significant Determination Process (IMC 0609, “Significance Determination Process”); and (4) Inspection Reports (IMC 0612). One out of every four feedback forms received was for IMC 0609. The next three documents with the most forms, with each chapter receiving about 10% of all feedback forms, were IMC 0305, IMC 0608, and IMC 0612.

The concentration of feedback forms in certain topical areas is consistent with the staff’s current improvement efforts in the reactor oversight process. IMC 0612 was issued in April 2002 to provide improved clarification to documentation of inspection findings; the SDP improvement program is on-going and the staff is currently working with the industry in developing the MSPI to replace the safety system unavailability performance indicator.

The number of feedback forms received in CY 2002 (112 forms) was nearly identical to the number of feedback forms received during CY 2001 (118 forms). Although this metric was met based on the slightly declining trend, the concentration of feedback forms in selected program areas indicated that there needs to be further improvement in these areas for CY 2003.

IP-4 Completion of Baseline Inspection Program

Definition: Annual completion of baseline inspection program.

Criteria: Defined as per IMC 2515, "Light-Water Reactor Inspection Program - Operations Phase." Regions report any non-completions at the end of each annual inspection cycle.

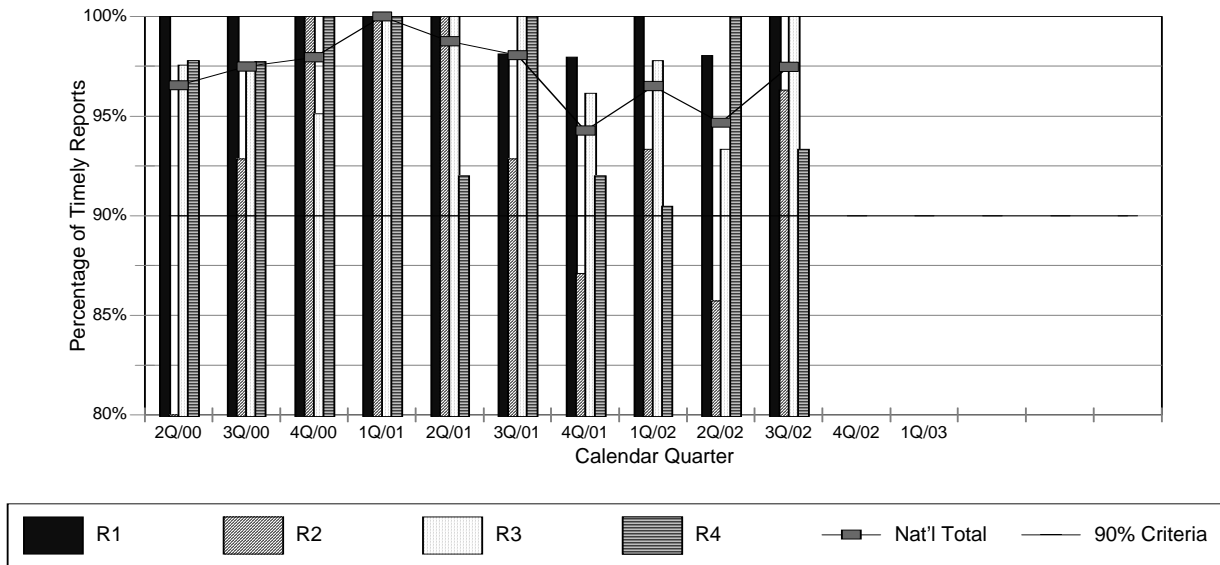
Analysis: The baseline inspection program was completed during ROP cycle 3 (CY 2002) in all regions. Davis-Besse was not included in this analysis since the baseline inspection program was replaced by inspections governed by the IMC 0350, "Oversight of Operating Reactor Facilities in an Extended Shutdown as a Result of Significant Performance Problems," process. All baseline inspections of annual periodicity were completed in CY 2002. In addition, all biennial inspections were completed at least once by the end of CY 2002. All triennial inspections were completed by March 31, 2003 (3 years from start of ROP), with one inspection in progress.

The inspection staff faced challenges during CY 2002. These challenges stemmed largely from a shortage of qualified inspectors and use of inspection resources to respond to unforeseen emerging events and external information demands. In response to these challenges, regional staff developed and implemented short-term coping strategies to complete the baseline inspections at all plants. Additionally, the staff has assessed inspection resource impact during CY 2003 and is working with the regions on potential short-term and long-term resolutions to this issue. A more detailed analysis of the challenges experienced with completing the baseline inspection program in CY 2002 and plans to address these concerns in CY 2003 is included in Attachment 7 to this paper.

IP-5 Inspection Reports are Timely

Definition: Obtain RPS data on the total number of reports issued and the number issued within timeliness goals (45 days for team and consolidated reports, 30 days for others).

Criteria: Expect 90 percent of inspection reports to be issued within program's timeliness goals.

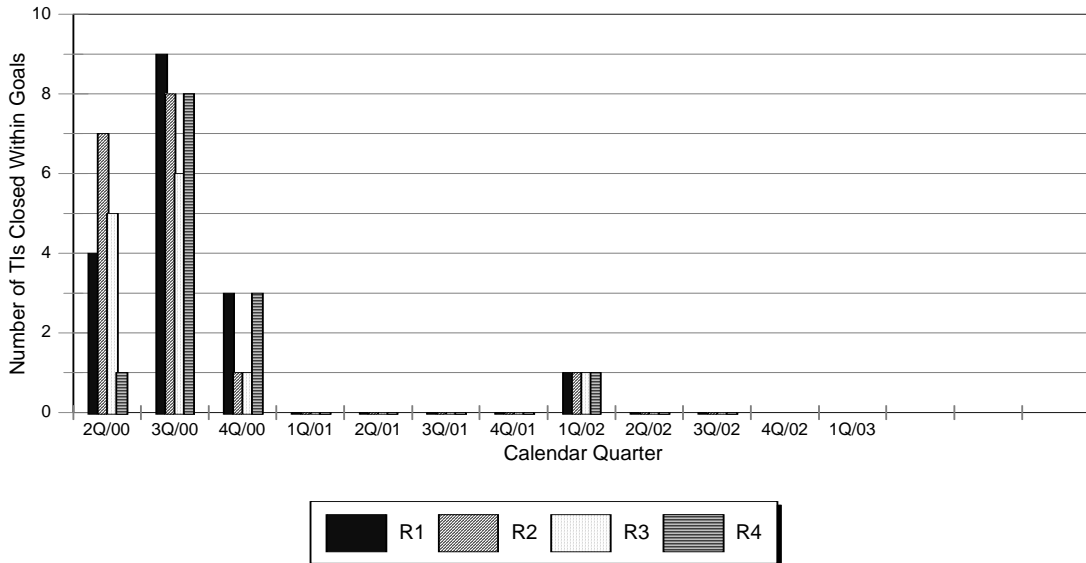


Analysis: A total of 412 inspection reports were issued through the third quarter of 2002. Overall as a program, 96 percent of all issued inspection reports were timely. Additionally, all regions met the inspection report timeliness goals during the calendar year 2002.

IP-6 Temporary Instructions (TIs) are Completed Timely

Definition: Audit the time to complete TIs by region. Compare the completion status in RPS to TI requirements. Report by region the number of TIs closed within goals.

Criteria: Expect all TIs to be completed within TI requirements.



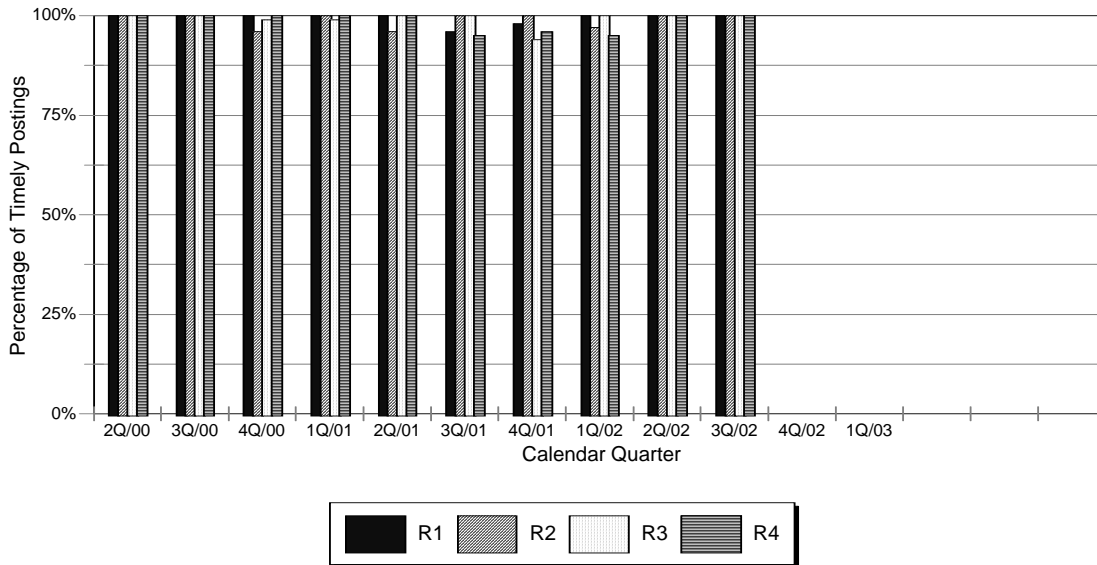
Analysis: TI 2515/144, "Performance Indicator Data Collecting and Reporting Process Review," was completed during the first quarter of 2002. All regions met the timeliness goals for completing TIs during calendar year 2002. In addition, there are currently seven TIs open.

IP-7 Public Communication Is Timely

Definition: The Inspection Program Branch (IIPB) within NRR posts inspection reports to the NRC’s external (public) Web site within ROP timeliness goals using an electronic version of inspection reports entered into Agency Document Access and Management System (ADAMS) by the regions. IIPB also posts entries from the Plant Issues Matrix (PIM) to the NRC’s public Web site using data entered into RPS by the regions. In addition, IIPB records the number of inspection reports not available in ADAMS and the number of PIM entries not updated in RPS, as well as the number of inspection reports and PIMs that are not posted to the NRC’s public Web site within goals.

Within 5 weeks of the end of each quarter, IIPB posts issued inspection reports from the previous quarter, using the electronic version in ADAMS, and the associated PIM entries from RPS to the NRC’s public Web site. Within 9 weeks of the end of each quarter, IIPB posts additional inspection reports and PIM entries for those not yet issued by the 5-week posting to include all findings from the previous quarter.

Criteria: Expect few untimely postings of PIMs or inspection reports, with a declining or stable trend in untimely postings.

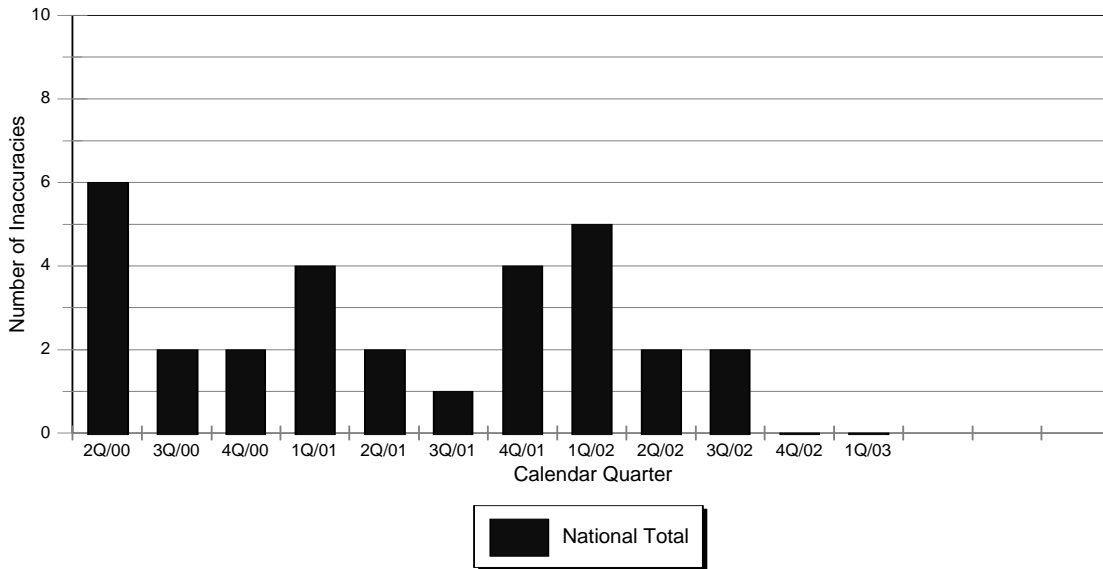


Analysis: There have been a few scattered untimely postings of inspection reports and/or inspection findings to the external Web site since the inception of the ROP. However, the percentage of timely postings has consistently been at or very near 100% for each quarter, with a stable trend in untimely postings.

IP-8 Public Communication Is Accurate

Definition: Each calendar quarter, sample information on the NRC’s external (public) Web site and count the number of times and reasons for regions changing PIMs or inspection reports (i.e., inaccuracy, new information).

Criteria: Track and trend.



Analysis: Inaccurate postings of PIM entries and inspection reports on the web were reasonably low during the past year, with an improving trend. The regions have issued more than 400 inspection reports and made more than 600 PIM entries during the year, indicating that the web accuracy percentage for inspection information is very high.

IP-9 Analysis of Inspection Hours

Definition: Collect and analyze RPS data (number of samples, regular hours, overtime hours) for each inspection procedure (including Plant Status). Collect preparation and documentation time.

Criteria:

- (1) Expect no significant deviations (less than 10% per procedure across all plants in region), and explore reasons for such deviations.
- (2) Track and trend overtime for the baseline inspection program and the underlying reasons, and use first year data to establish a baseline.
- (3) Track and trend preparation, documentation, travel, and communication times to establish a baseline, and assess the effects on budgeted resources.

Analysis: Total staff efforts to complete baseline activities decreased by about 10% as compared to the first two years of ROP implementation. Resources expended in direct inspection effort and inspection preparation and documentation appear to remain constant. Overall, there was a decrease of 10% in resources expended to complete the ROP in fiscal year 2002 as compared to the previous two ROP periods. For a more detailed analysis of ROP resources, see Attachment 7 to this paper.

IP-10 Survey of ROP Users

Definition: Survey inspectors and other NRC personnel implementing the ROP, asking whether the inspection program covers areas that are important to safety.

Criteria: Trend average level of agreement.

Analysis: About 70 percent of those surveyed in November 2002 agreed that the baseline inspection program appropriately inspected for and identified risk-significant issues. More than 80 percent of the survey respondents indicated that inspection reports were communicated accurately and in a timely fashion.

Also, more than 70 percent of those surveyed indicated that the inspection procedures were (1) adequate to address intended cornerstone attributes, (2) clearly written, (3) written to place sufficient emphasis on planning, and (4) conducted at an appropriate frequency. About 75 percent of those surveyed agreed that the inspection procedure adequately sampled risk important aspects of each inspectable area.

More than 60 percent of those surveyed agreed that the baseline inspection program report format adequately communicated relevant information to the licensee and to the NRC internal stakeholders. About 63 percent of those surveyed agreed that the report format adequately communicated relevant information to the public.

The November 2002 survey results regarding the inspection program were generally favorable and were comparable to those from the March 2001 survey; therefore, this metric was met.

IP-11 Survey of Inspection Report Usefulness

Definition: Survey external stakeholders, asking about the usefulness of inspection reports. This question will be included in the Federal Register notice.

Criteria: Trend average level of agreement.

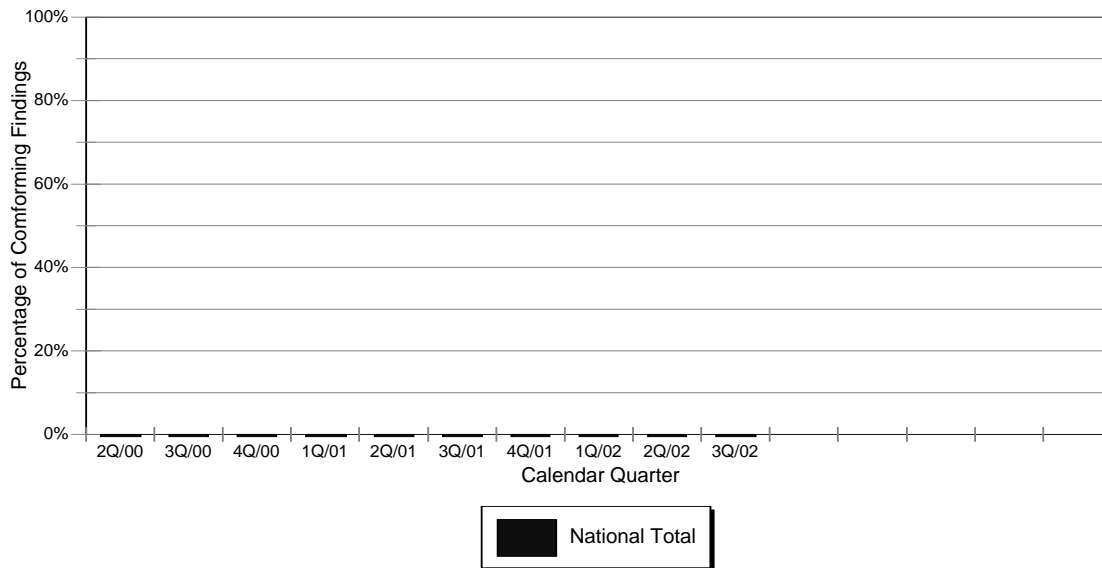
Analysis: The majority of those who provided feedback to the question (10 out of the 11 responses) on whether the information in the inspection reports were useful to them responded favorably. Two responders indicated that the inspection report would be more useful if it contained more information to allow trending or evaluation of less significant events. Six responders did not provide feedback on the quality of information in the inspection reports.

This metric was met based on a similar level of positive response when compared to the previous survey.

SDP-1 The SDP Results Are Predictable and Repeatable and Focus Stakeholder Attention on Significant Safety Issues

Definition: Quarterly audit of a representative sample of reported inspection findings against the standard criteria set forth in IMC 0609. Findings should contain adequate detail to enable an independent auditor to trace through the available documentation and reach the same significance color characterization.

Criteria: The target goal is at least 90% are determined to be predictable and repeatable. Any SDP outcomes determined to be non-conservative will be evaluated and appropriate programmatic changes will be implemented.



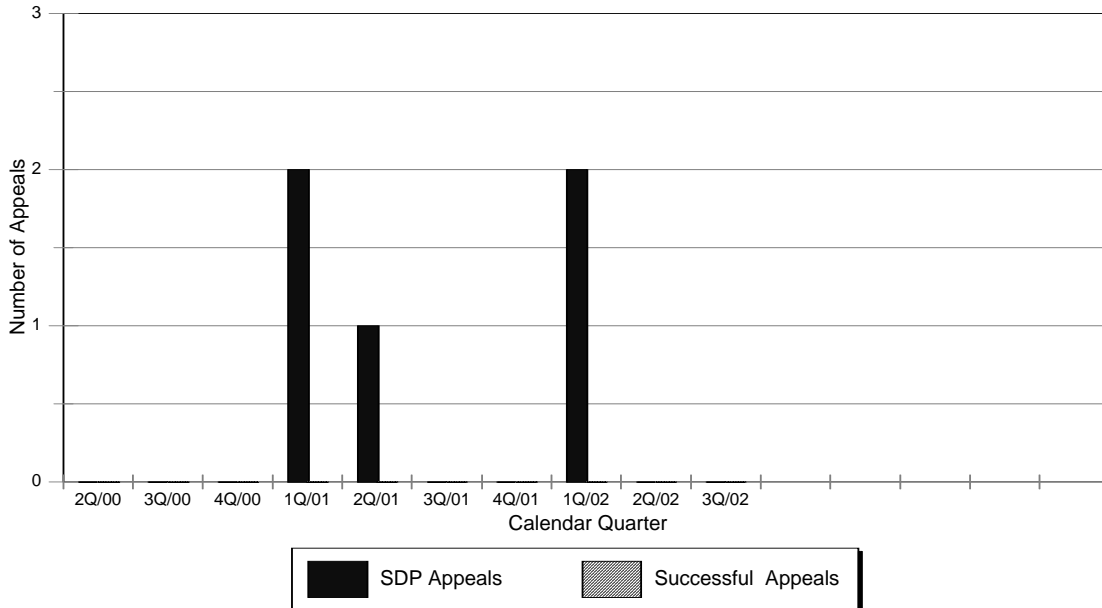
Comments: The staff issued IMC 0612 on April 29, 2002, to improve the program guidance on documentation of inspection findings. After a brief training period, all regions implemented the new requirements of IMC 0612 in July of 2002. To allow inspectors and regional management to become more familiar with and implement the new requirements of IMC 0612 and time to issue a sample inspection report for regional use, the staff agreed to allow the first set of inspection report audits under IMC 0612 be conducted by regional personnel. The staff will commence auditing the inspection reports in CY 2003.

Analysis: No data was available in CY 2002 due to program transition.

SDP-2 SDP Outcome Is Risk-Informed and Accepted by Stakeholders

Definition: Track the total number of appeals of final SDP results reported quarterly by the regions.

Criteria: Expect zero appeals of SDP significance that result in a final determination being overturned across all regions.



Analysis: During FY 2002, there were two appeals of final SDP outcomes. The appeals were submitted during Q1/2002 and involved a White Public Radiation Safety SDP finding at Comanche Peak and a Yellow Emergency Preparedness SDP finding at Columbia Generating Station. Both SDP outcomes were upheld following the appeal process.

Performance during this assessment period met program expectations.

SDP-3 Inspection Staff Is Proficient and Finds Value in Using the SDP

Definition: Survey internal stakeholders using specific quantitative survey questions that focus on training, effectiveness, and efficiency.

Criteria: Expect either a stable or an increasingly positive perception of the SDP process over time.

Analysis: Of the internal stakeholders participating in the survey that addressed these specific questions, 71% of the respondents agreed that the SDP focuses NRC attention on safety-significant issues, 73% felt the SDP provides an effective basis for communicating findings to the licensee, and 60% said the SDP provides an effective basis for communicating findings to the public. Finally, 61% felt the SDP provides for consistent results. On the negative side, 33% felt that SDP training was effective, 20% said the reactor safety SDP was easy to use, and 26% said the non-reactor safety SDPs were easy to use. Approximately 33% of those surveyed felt the program guidance documents were clear, while 26% agreed that the resource expenditures were appropriate. Overall, the survey results indicate that the staff believes the SDP is effective in meeting program objectives, but continue to express skepticism regarding their proficiency in completing phase 2 SDP evaluations. These results are similar to those noted after the March 2001 internal survey.

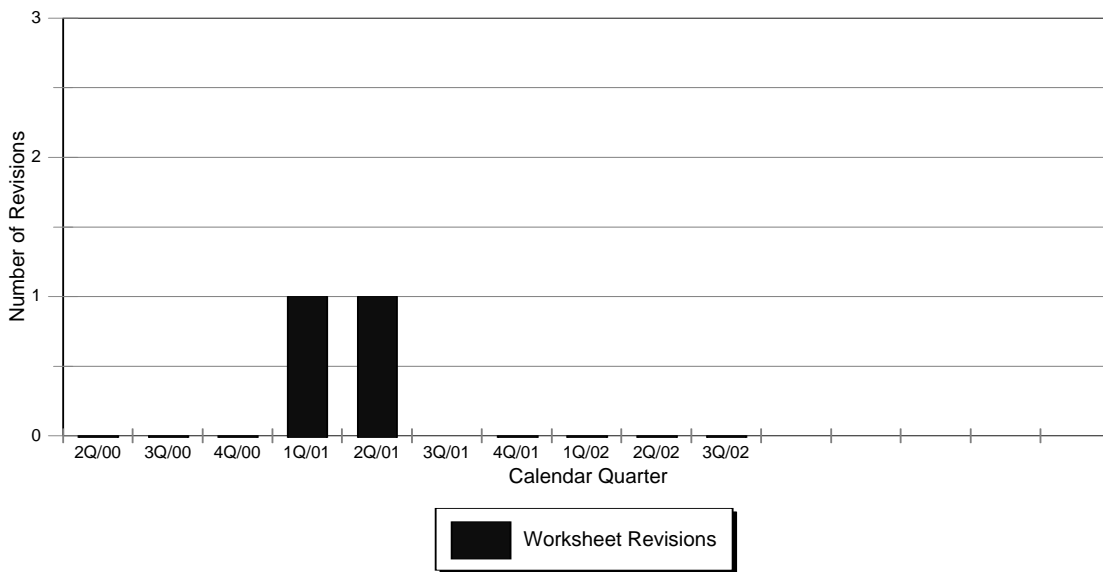
Improvement strategies noted in the SDP Improvement Initiative that address development of improved SDP training and SDP tools are focused on improving inspector proficiency and confidence in completing SDP phase 2 evaluations. In addition, in response to internal stakeholder comments, the staff initiated an SDP Task Group. The staff will be evaluating and implementing the SDP Task Group recommendations to help develop improvements in this area.

The inspection staff generally agreed that the SDP was useful, but still noted concerns with staff proficiency in using the SDP. Although the staff could conclude that the specified criteria was met based on similar results found from the March 2001 survey, the low percentages of individuals who felt that the SDP training was effective and that the SDPs were easy to use does not meet the staff's expectations. Therefore, this metric has not been met.

SDP-4 SDP Tools for Evaluating Inspection Findings Reflect Current Plant Design and Licensee Operating Practices

Definition: Monitor substantive revisions made to the risk-informed inspection notebooks due to non-conservative technical flaws by tracking the number of phase 2 inspection notebooks that are issued for use and subsequently withdrawn following onsite benchmarking activities.

Criteria: The target goal is zero notebook retractions due to non-conservative technical flaws.



Analysis: The staff recognized the need to benchmark the inspection notebooks, and has stepped up an aggressive schedule to complete the benchmarking of all Phase 2 notebooks by the end of FY2003. The risk-informed inspection notebooks for 48 reactor facilities have been validated by benchmarking, which included comparing the notebooks against licensee-developed risk models using similar assumptions. No (revision 1) notebooks have been retracted or returned to Brookhaven National Laboratories for immediate revision to limit potentially non-conservative outcomes during the assessment period. Risk notebooks retracted during the previous assessment period (FY 2001) for Calvert Cliffs and South Texas Project have since been benchmarked and reissued as revision 1.

Performance during this assessment period met program expectations.

SDP-5 Results of the Same Color Are Perceived by the Public to Translate to the Same Level of Significance for All Cornerstones

Definition: Publish a *Federal Register* notice to survey external stakeholders using specific questions asking for examples of where the SDP-determined significance of findings does not appear to be consistent across ROP cornerstones.

Criteria: Expect stable or increasingly positive perception of the SDP over time.

Analysis: External stakeholder survey results indicated that industry respondents did not believe that same color findings are consistent across all ROP cornerstones. Industry respondents felt that the reactor safety cornerstones were consistent, but were dissatisfied with SDP results for Emergency Preparedness, Radiation Safety, and Physical Protection. The impression was that these SDPs were not risk-informed, but “a deterministic escalation for various types of regulatory noncompliance,” and in general, that these SDPs were subjective in nature. The results also indicated that non-green thresholds for these SDPs may overstate the significance of findings.

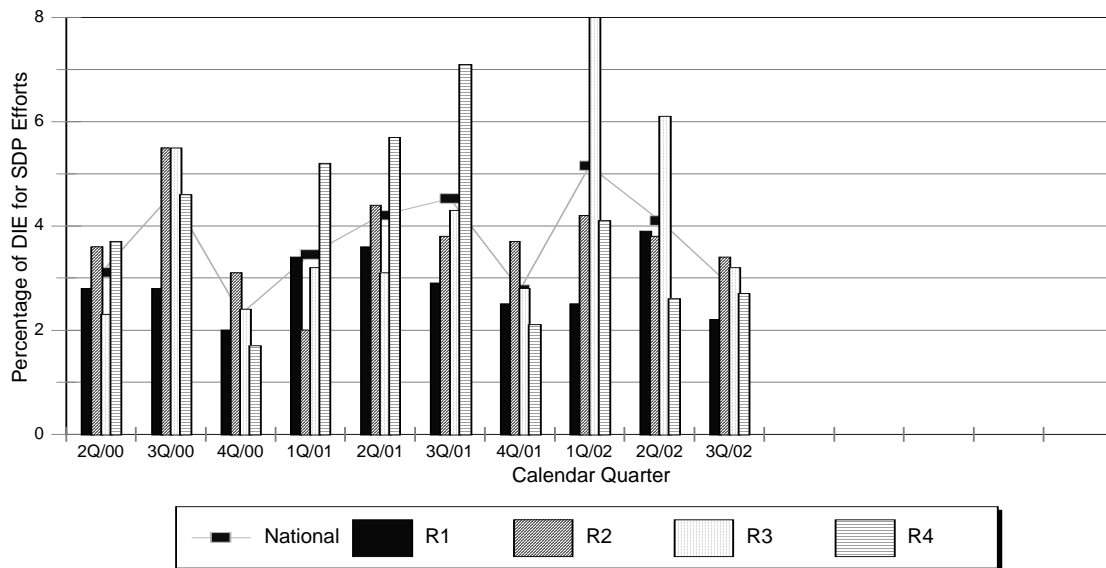
Although no specific comments were made regarding the significance of findings across cornerstones, citizens groups insisted that downgrades of preliminary SDP findings are a result of industry lobbying efforts which attempt to water down findings and undermine the legitimacy of the ROP.

Performance during this assessment period did not meet program expectations.

SDP-6 The Resources (Direct Charges and Support Activities) Expended Are Appropriate to the Benefit (Significance of Issues Identified)

Definition: Track the percentage of total inspection resource expenditures attributed to SDP activities. Calculate the effort expended by the regions in completing SDP evaluations as a percentage of the total regional direct inspection effort. Use RPS codes for SDP processing activities.

Criteria: Total SDP expenditures should not exceed 10 percent of the total regional direct inspection effort (DIE) with a stable or decreasing trend over time.



Analysis: Although the reported regional expenditures associated with SDP evaluations remain below the target goal, the average SDP evaluation time increased substantially during the Q1/2002. The staff evaluated potential causal factors and determined that the increase in SDP resource expenditures was due, in part, to the reviews of complex engineering issues identified at D.C. Cook, Point Beach, and Davis-Besse. The resource expenditures trended downward during subsequent quarters. Lower resource expenditures in Q3/2002 can be attributed to fewer inspection findings reported during that quarter.

Performance during this assessment period met program expectations.

SDP-7 Appropriateness of Regulatory Impact From the SDP

Definition: Monitor the trend of regulatory impact forms that are critical of the SDP and assessment processes.

Criteria: Stable or decreasing trend over time.

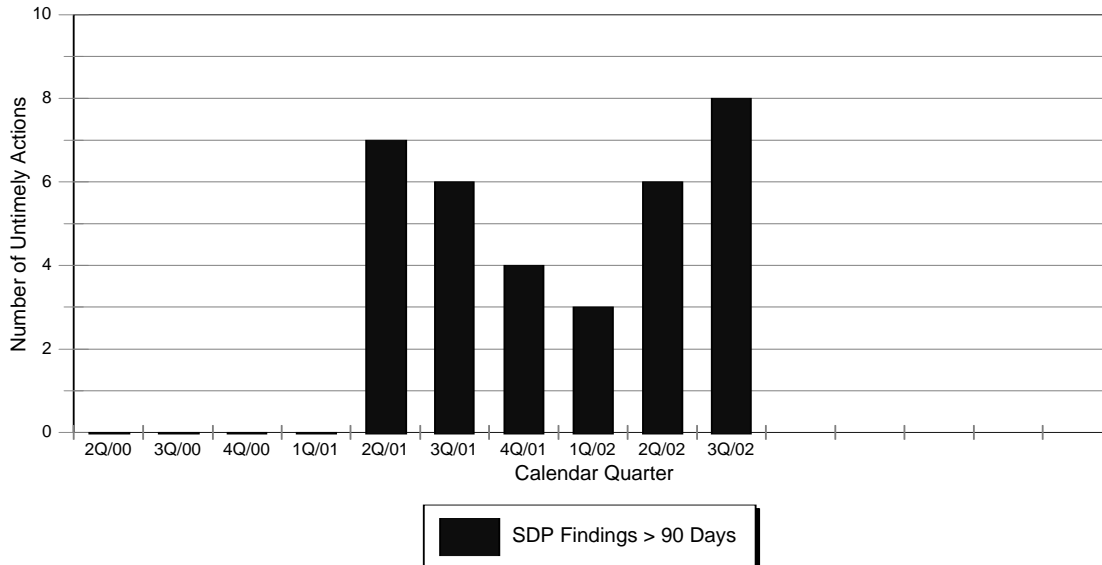
Comment: This metric is not available as the applicable staff was redirected to support activities in the Office of Nuclear Security and Incident Response. While the process of using regional managers to solicit feedback from power-reactor licensees is functioning normally, the annual evaluation of this feedback was postponed a year. The next evaluation of licensee feedback will span two years and will be conducted in late CY 2003.

SDP-8 Final Significance Determinations Are Timely

Definition: Conduct a quarterly audit of RPS data to identify the total number of inspection items that have been under review for more than 90 days since:

- (1) the date of initial licensee notification of the preliminary significance in an inspection report, or
- (2) the date the item was formally transmitted to an NRR technical branch for SDP assistance, or
- (3) the item was otherwise documented in an inspection report as an unresolved item pending completion of a significance determination and not counted in either of the above categories.

Criteria: All SDP results that are counted per the criteria above should be finalized within 90 days.



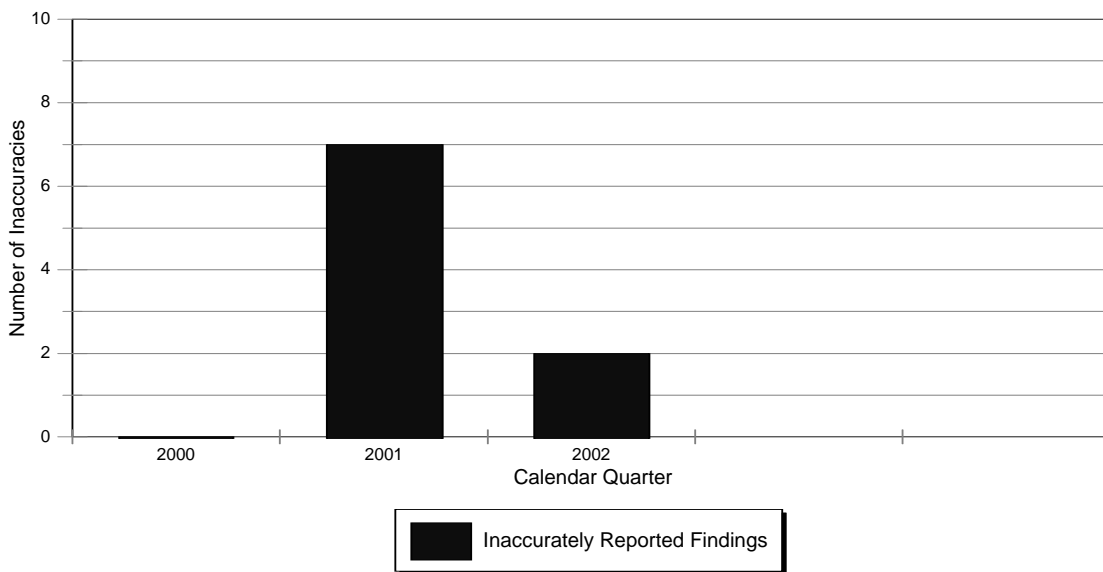
Analysis: In response to Commission direction, the staff has adjusted the criteria for measuring SDP timeliness to monitor for final issuance of SDP findings within 120 days of the initial exit meeting and 90 days of official licensee notification in an inspection report. This adjustment to the criteria has been included in the SDP timeliness strategies that are currently under review by senior NRC management. During this assessment period several instances of late significance determinations were identified. Performance in this area is a continuing challenge due to the complexity of some of the issues, and is being addressed by the SDP Improvement Initiative and the SDP Active Issues Matrix which was developed to focus regional and headquarters management attention on prompt resolution of more risk significant issues.

Performance during this assessment period did not meet program expectations.

SDP-9 SDP Results Are Communicated Accurately to the Public

Definition: Each calendar quarter, track the number of inspection findings that are inaccurately communicated to the public (color of findings is inaccurately reported), by auditing the inspection findings summary information available on the NRC Web. The detailed review will include item type, significance characterization, enforcement action status, and text descriptions of greater-than-green inspection findings prior to release to external stakeholders.

Criteria: The target goal is zero inaccuracies. All inaccuracies must be addressed.



Analysis: During the current assessment cycle, two instances were identified in which the status of documented inspection findings reported on the NRC’s external web site was unclear (i.e., preliminary vs. final) when looking at assessment information developed from the reported inspection finding data. In one instance the final determination was not posted to the web in a timely manner, and the plant was not reflected in the appropriate Action Matrix column for several days. In the other instance, the issue was double counted and placed in the wrong quarter on the web site. Quarterly audits identified both issues and the information on the web was corrected immediately.

Performance in this area has improved but is still not meeting the established criteria of zero inaccuracies.

AS-1 Subjective Judgment Is Minimized and Is Not a Central Feature of the Process. Actions Are Determined by Quantifiable Assessment Inputs (Examine PIs and SDP Results)

Definition: Audit all assessment-related letters and count the number of deviations from the Action Matrix.

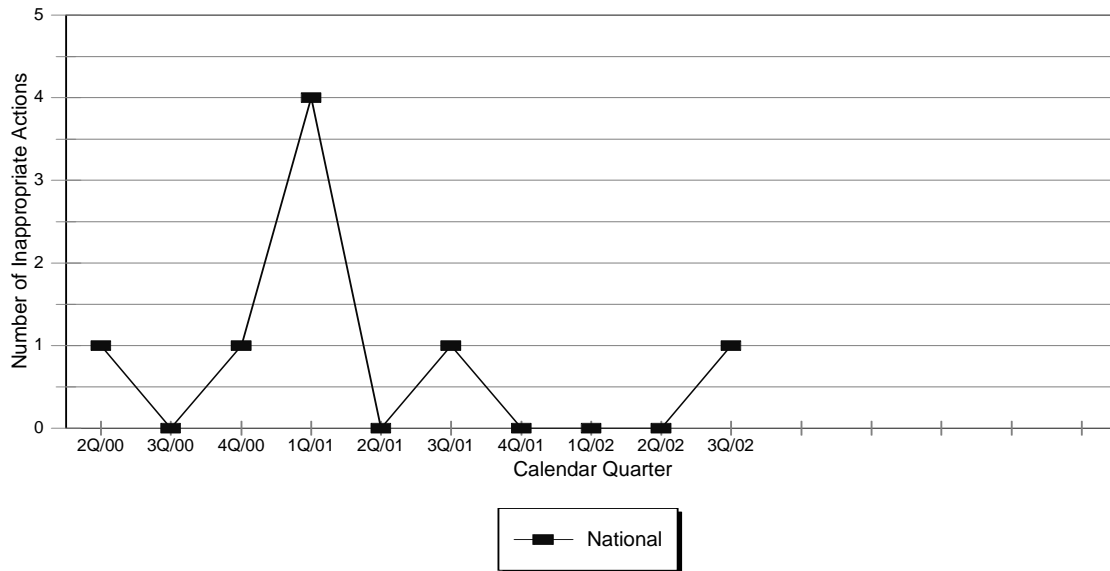
Criteria: Expect few deviations, with a declining trend.

Analysis: There was one deviation from the Action Matrix for Oconee Unit 1. By letter dated August 26, 2002, the NRC approved a deviation from actions required by IMC 0305 for a plant in the multiple/repetitive degraded cornerstone of the Action Matrix. This was the first Action Matrix deviation that has been approved since the beginning of the ROP; accordingly, this metric is considered to have been met.

AS-2 The Program Is Well-Defined Enough to Be Consistently Implemented

Definition: Audit all assessment letters and count the number of significant departures from requirements in IMCs 0305 and 0350. Timeliness goals are counted in metric AS-5.

Criteria: Few departures, steady or declining trend.

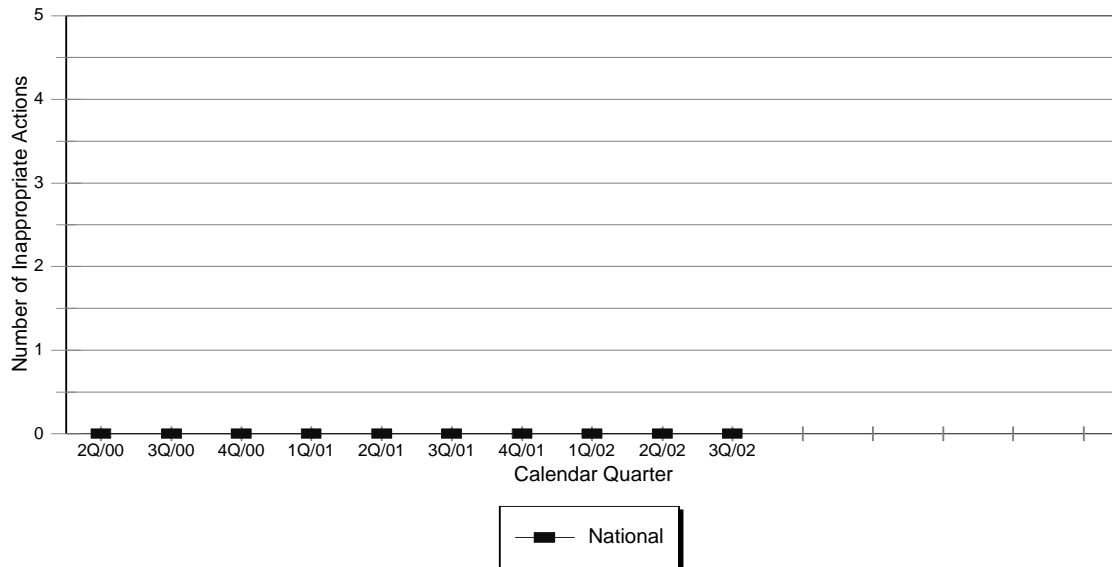


Analysis: One mid-cycle letter for a plant in the degraded cornerstone column of the Action Matrix in 3Q/2002 was signed by the regional division director instead of the regional administrator. This appears to be an isolated occurrence and meets the criteria of few departures with a steady or declining trend.

AS-3 Actions Taken Are Commensurate With the Risk of the Issue and Overall Plant Risk

Definition: Review actions taken for greater-than-green inspection findings and PIs. Track the number of actions (or lack of actions) taken by the regions that are not appropriate for the significance of the issues and are not consistent with the Action Matrix.

Criteria: Expect few departures, with a steady or declining trend.



Analysis: All actions taken by the regional offices were consistent with the Action Matrix during the period of October 2001- September 2002.

AS-4 The Number And Scope of Additional Actions Recommended as a Result of The Agency Action Review Meeting (AARM) Beyond Those Actions Already Taken Are Limited

Definition: Review the results of the AARM.

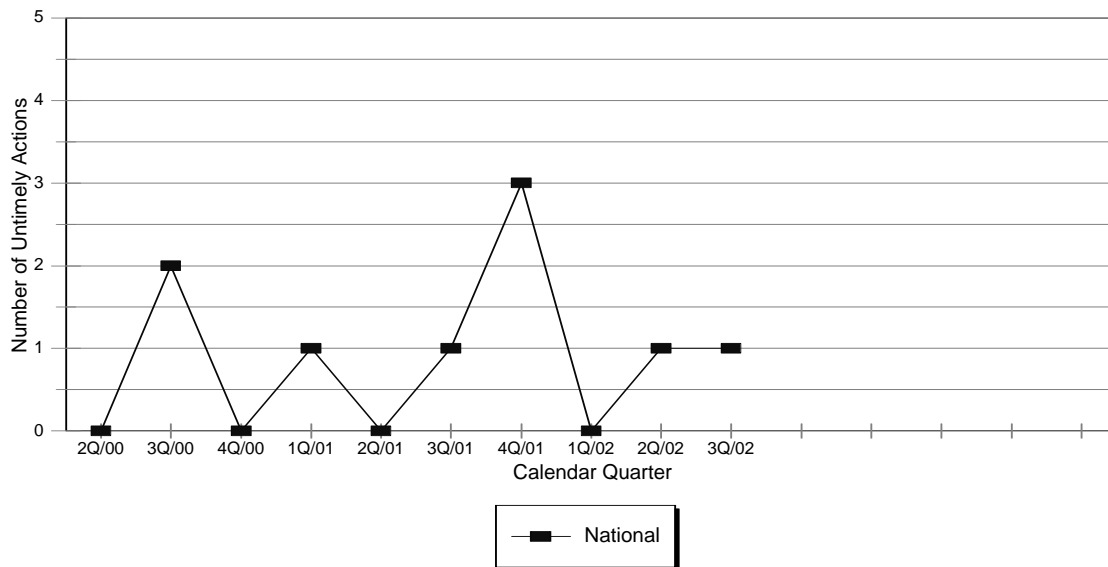
Criteria: The AARM should recommend few additional actions, with a steady or declining trend from the first-year benchmark.

Analysis: The AARM was held on April 9 - 11, 2002, in Annapolis, Maryland. The participants confirmed the appropriateness of agency actions for Cooper and Indian Point 2. The participants did not recommend any additional actions beyond those already taken or planned. The next Agency Action Review Meeting is scheduled for April 22 - 23, 2003.

AS-5 Assessment Program Results (Assessment Reviews, Assessment Letters and Public Meetings) Are Completed in a Timely Manner

Definition: Track the number of instances in which timeliness goals established in IMC 0305 were not met. Collect timeliness data for the conduct of quarterly reviews (within 5 weeks of the end of quarter); mid-cycle and end-of-cycle reviews (within 6 weeks of the end of quarter); issuance of assessment letters (within 2 weeks of the quarterly review and 3 weeks of the mid-cycle and end-of-cycle reviews); assessment followup letters (on or before the next quarterly review); and public meetings (within 16 weeks of the end of the assessment period).

Criteria: Expect few instances in which timeliness goals were not met, with a steady or declining trend from the first-year benchmark.



Analysis:

3Q/2002: All sixty-six mid-cycle review meetings and associated mid-cycle letters were completed within timeliness goals. One assessment follow-up letter was not issued within timeliness goals.

2Q/2002: Two out of three assessment follow-up letters were issued within timeliness goals.

1Q/2002: All sixty-six end-of-cycle meetings, annual assessment letters, and annual public meetings were completed within timeliness goals.

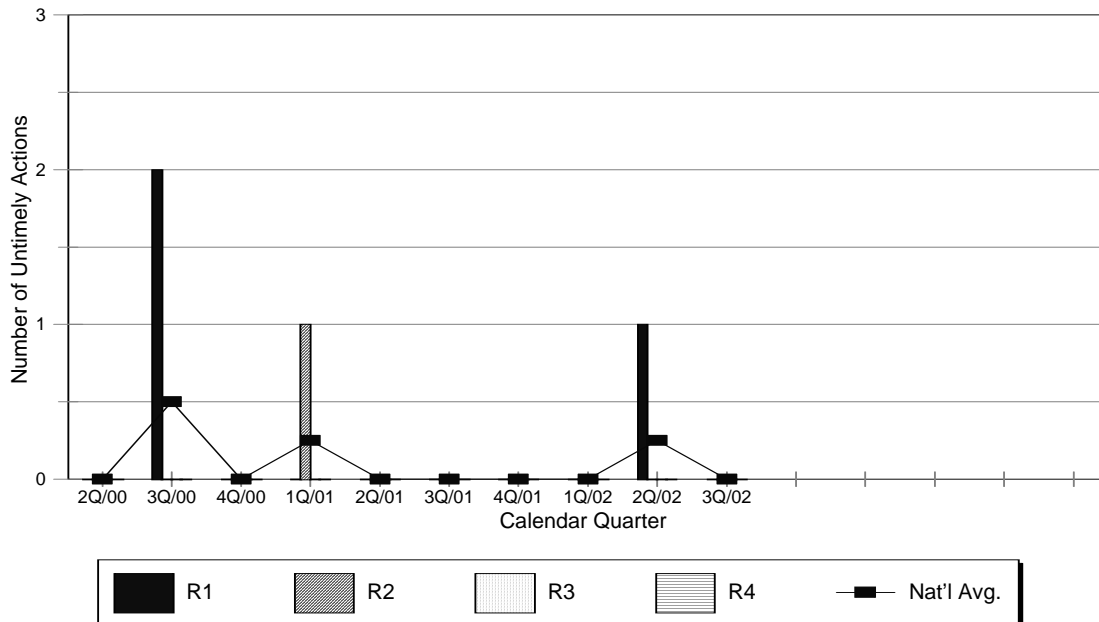
4Q/2001: Ten of the thirteen assessment follow-up letters were issued within timeliness goals.

Timeliness goals for completion of assessment program activities were achieved at a rate of approximately 99%. Therefore, this performance metric was met.

AS-6 The Web Posting and Availability Via ADAMS of Assessment Letters Is Timely

Definition: Review the posting of letters to the NRC’s external Web site and availability in ADAMS and compare to the timeliness goals. Record the number of letters not available in ADAMS and number of letters not posted to the Web site within goals.

Criteria: IIPB posts assessment letters to the NRC’s external Web site using the electronic version in ADAMS within 10 weeks of the end of mid-cycle and end-of-cycle assessment periods and within 8 weeks of the end of intervening quarters.



Analysis:

Q3/2002: All 66 mid-cycle letters were posted to the web within timeliness guidelines.

Q2/2002: One assessment follow-up letter was not posted to the web within timeliness goals.

Q1/2002: All 66 annual assessment letters were posted to the web within timeliness guidelines.

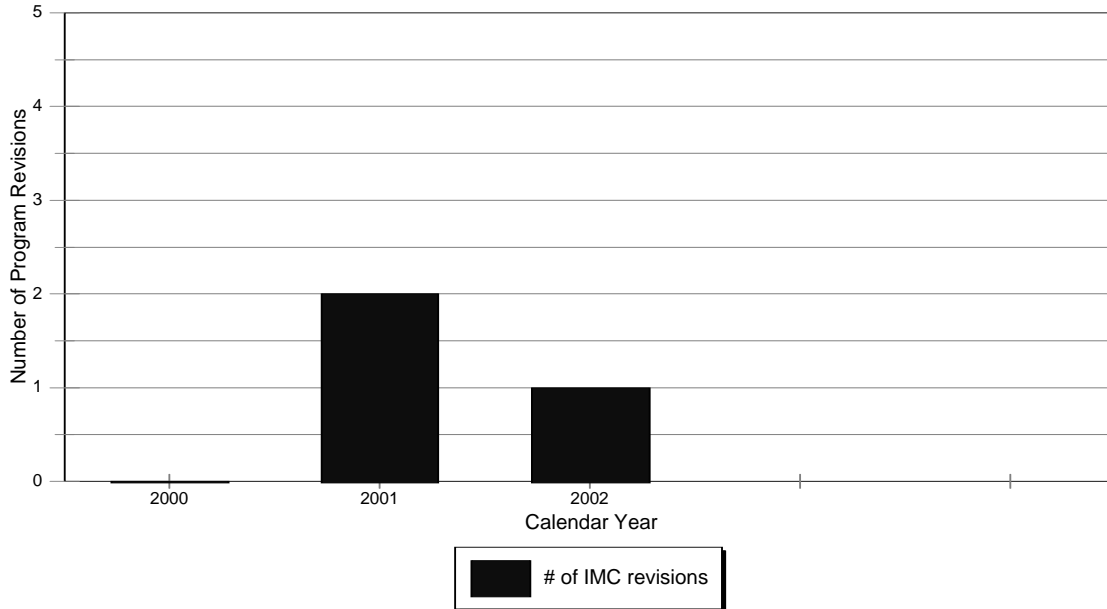
Q4/2001: Most assessment letters were not posted to the web due to the terrorist attacks on September 11, 2001. One noted exception was two assessment follow-up letters that were posted on September 2, 2001.

With the exception of the decision to hold off posting ROP and other information due to the terrorist attacks on September 11, 2001, the timeliness goals for posting assessment letters to the NRC’s external Web site were met.

AS-7 Assessment Program Procedures Are Stable Enough To Be Perceived as Predictable

Definition: Count the number of revisions to IMCs 0305 and 0350.

Criteria: Expect few revisions, with a steady or declining trend from the first-year benchmark.

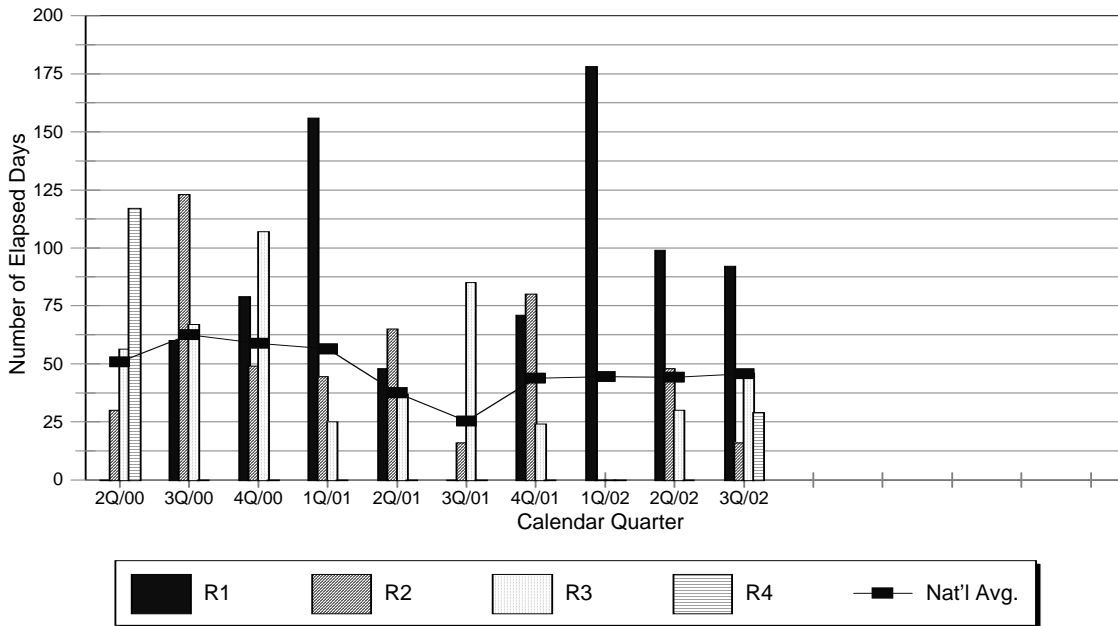


Analysis: During CY 2002, there was one revision to IMC 0305 which was issued on February 11, 2002. Another revision to IMC 0305 was issued in February 2003. A revision to IMC 0350 is also planned for CY 2003 to incorporate insights and lessons learned from the Davis-Besse event. Therefore, this performance metric is considered to have been met.

AS-8 The NRC's Response to Performance Issues Is Timely

Definition: Count the number of days between issuance of an assessment letter discussing an issue of more than very low safety significance and completion of the supplemental inspection (by exit meeting date, not issuance of the inspection report).

Criteria: Expect a steady or declining trend when compared to the benchmarking data (first few years of the ROP).



Analysis: Baseline data for this metric are still being collected. However, data collected to date does not indicate a negative short-term trend regarding the elapsed time between the issuance of an assessment letter and the completion of the corresponding supplemental inspection.

AS-9 The Agency Takes Appropriate Actions To Address Performance Issues for Licensees Outside of the Licensee Response Column of the Action Matrix

Definition: Solicit feedback on the appropriateness of regulatory attention given to licensees with performance problems via a survey question to both internal and external stakeholders.

Criteria: Expect steady or improved perception of appropriateness of actions as compared to the first-year benchmark.

Analysis:

Internal survey

Listed below are the staff's responses to the following statement: "The agency takes appropriate actions to address performance issues for those licensee outside of the Licensee Response Column of the Action Matrix".

Strongly agree: 7.8%
Agree: 64.7%
Disagree: 12.1%
Strongly disagree: 6.0%
Unable to answer: 9.5%

Some of the staff's additional comments on this question indicated a level of concern with the ability of the NRC to detect declining performance in a timely manner. The vessel head degradation at Davis-Besse was discussed as an example where declining performance was not detected early. Some of the staff also expressed concern with the lack of prominence that cross-cutting issues receive in the assessment program.

This question was not specifically asked during the March 2001 internal survey, so we have no basis to compare to previous survey results. However, the respondents generally agreed that the agency takes appropriate actions to address performance issues for those licensees outside of the licensee response column of the action matrix.

External survey

The public interest groups were generally negative regarding the NRC's actions to address performance issues for plants outside of the Licensee Response Column of the Action Matrix. Two of these participants stated that the agency is not taking actions mandated by the Action Matrix but merely changing the color of inspection findings to justify the desired response in the Action Matrix. The industry and two States (Illinois and Pennsylvania) were positive regarding the NRC's actions to address performance issues for plants outside of the Licensee Response Column of the Action Matrix.

This metric meets its criteria with a stable perception when compared to previous survey results.

AS-10 Information Contained in Assessment Reports Is Relevant, Useful, and Written in Plain Language

Definition: Perform surveys to determine internal and external stakeholder views on assessment reports.

Criteria: Steady or improved perception of the relevance, usefulness, and understandability of assessment reports as compared to the first year benchmark.

Analysis:

Internal survey

Listed below is the staff's average response to the statements concerning whether the information contained in the assessment letters is relevant, useful, and written in plain language.

Strongly agree: 8.1%
Agree: 60.5%
Disagree: 19.4%
Strongly disagree: 7.2%
Unable to answer: 4.8%

This question was not specifically asked during the March 2001 internal survey, so we have no direct basis to compare to previous survey results. However, the respondents generally agreed that the information contained in the assessment letters is relevant, useful, and written in plain language.

External survey

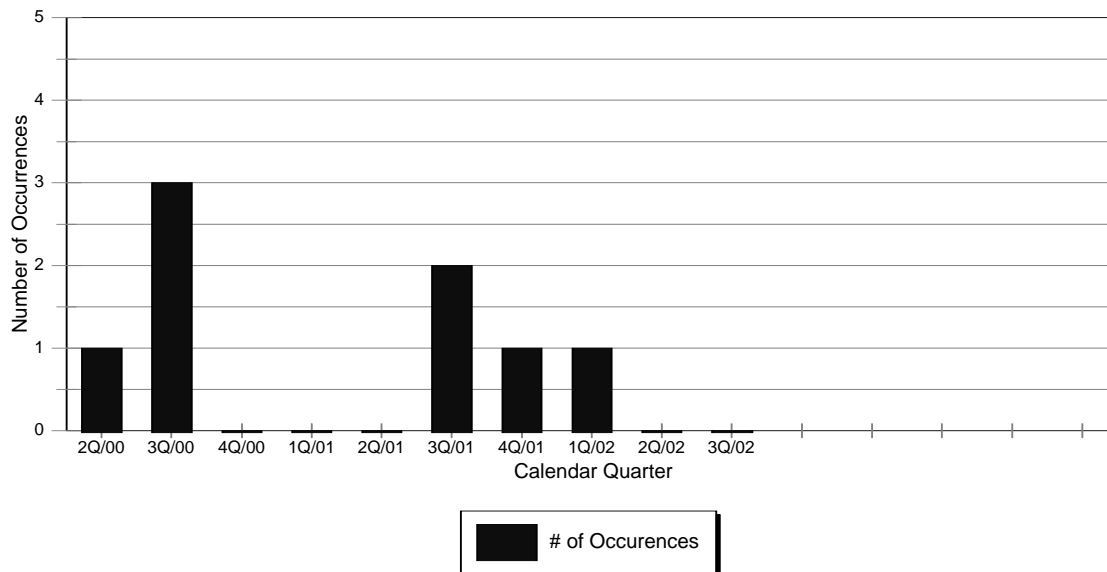
Feedback from public interest groups was mixed. One public interest group responded positively but two others added that the assessment letters were of little value. The industry responded positively but added that the annual public meetings should be used as an opportunity for more public outreach. One industry participant added that the annual public meetings should be eliminated for plants that have all green performance indicators and inspection findings. Responses from the State regulators were generally positive and recognized an improvement in the assessment report quality over the last few years. One State responded that there were insufficient details in the assessment letters to allow trending.

This metric meets its criteria with a stable perception when compared to previous survey results.

AS-11 Degradations in Plant Performance, as Measured in the Action Matrix, Are Gradual and Allow Adequate Agency Engagement of the Licensees

Definition: Track the number of instances each quarter in which plants move more than one column to the right in the Action Matrix (as indicated on the Action Matrix Summary).

Criteria: Expect few instances in which plant performance causes a plant to move more than one column to the right in the Action Matrix. Provide a qualitative explanation of each instance in which this occurs. Expect a steady or declining trend from the first year benchmark.



Analysis: During the period of October 2001 - September 2002, there were two plants that moved more than one column to the right in the Action Matrix in one quarter. In 4Q/2001, Columbia Generating Station moved from the licensee response column to the degraded cornerstone column of the Action Matrix due to a yellow finding in the emergency preparedness cornerstone. In 1Q/2002, Vermont Yankee moved from the licensee response column to the degraded cornerstone column of the Action Matrix due to a yellow finding in the physical protection cornerstone. The number of plants moving two or more columns to the right in the Action Matrix have been few and within the expected frequency.

O-1 Public Perceives the ROP to Be Predictable and Objective

Definition: Annually survey external stakeholders through a *Federal Register* notice asking if decisions are overly reliant on judgement, or not controlled by the process.

Criteria: Expect a stable or increasing positive perception over time.

Analysis: A majority of licensee respondents state the ROP is predictable and objective while non-licensee groups as a whole believe it is far too subjective and based on individual judgement or negotiation. The responses from licensees are similar to those from previous years and indicate a belief in the continuing effort to improve in this area. A few respondents cite different SDP outcomes as an example of the ROP being unpredictable and inconsistent. While there is not complete agreement on the objectivity of the process there is some agreement that it is moving in the right direction.

This metric meets its criteria with a stable and slightly increasing positive perception.

O-2 NRC Perceives the ROP to Be Predictable and Objective

Definition: Annually survey internal stakeholders, asking if decisions are overly reliant on judgement, or not controlled by the process.

Criteria: Expect stable or increasingly positive perception over time.

Analysis: Three internal survey questions addressed this metric. The percentage of NRC internal stakeholders agreeing that the ROP increases predictability remained positive and similar to the previous survey (75% in 2002 vs. 75% in 2001). Respondents agreeing that the ROP generally provides appropriate objectivity to the oversight process was reduced slightly (82% in 2002 vs. 85% in 2001) as was the belief that the ROP increases objectivity (76% in 2002 vs. 79% in 2001), but remained generally positive.

The data supporting this metric indicates a slightly decreasing positive perception when compared to the previous survey. Although this metric did not explicitly satisfy the established criteria, internal stakeholders generally agreed that the ROP is predictable and objective and the perception was relatively stable.

O-3 Public Perceives the ROP to Be Risk-informed

Definition: Annually survey external stakeholders through a *Federal Register* notice asking if ROP actions and outcomes are appropriately graded according to the significance of the issues at the plants.

Criteria: Expect stable or increasingly positive perception over time.

Analysis: Similar to previous surveys, overall a majority of respondents believe the ROP is risk informed. Respondents believe the ROP is a step increase in risk-informed regulation over the previous systems. The initiating events, mitigating systems, and barrier criteria cornerstones are specifically mentioned as being quite risk-informed. Numerous respondents do express concern for other cornerstones of the ROP that are less risk informed. Most negative comments to this question centered on the SDP and its perceived short comings.

This metric meets its criteria with a stable perception.

O-4 NRC Perceives the ROP to Be Risk-Informed

Definition: Annually survey internal stakeholders asking if ROP actions and outcomes are appropriately graded according to the significance of the issues at the plants.

Criteria: Expect stable or increasingly positive perception over time.

Analysis: Two internal survey questions addressed this metric. The percentage of NRC internal stakeholders agreeing that the ROP generally provides an effective risk-informed approach to oversight was reduced when compared to the previous survey (73% in 2002 vs. 82% in 2001), but remained generally positive. Compared to the previous process, respondents agreeing that the new ROP is more risk-informed was also reduced (91% in 2002 vs. 96% in 2001), but remained very positive.

The data supporting this metric indicates a slightly decreasing positive perception when compared to the previous survey. Although this metric did not explicitly satisfy the established criteria, internal stakeholders generally agreed that the ROP is appropriately risk-informed.

O-5 Public Perceives the ROP to Be Understandable

Definition: Annually survey external stakeholders through a *Federal Register* notice asking if they understand the process, procedures, and outputs, and if products are clear and written in plain English.

Criteria: Expect stable or increasingly positive perception over time.

Analysis: A significant majority of the stakeholders state that the ROP is understandable and that products are written in clear and plain English. Numerous stakeholders express reservations about the public's ability to understand the SDP and the color assignments for findings. As in the previous survey, the SDP is recognized to be the most complex portion of the ROP requiring some technical background for understanding.

This metric meets its criteria with a stable and increasing positive perception.

O-6 NRC Perceives the ROP to Be Understandable

Definition: Annually survey internal stakeholders asking if they understand the process, procedures, and outputs, and if products are clear and written in plain English.

Criteria: Expect stable or increasingly positive perception over time.

Analysis: Two internal survey questions addressed this metric. The percentage of NRC internal stakeholders agreeing that the ROP is understandable and written in plain English was stable when compared to the previous survey (87% in 2002 vs. 89% in 2001) and remained very positive. Note that no comparison could be made for the second internal survey question addressing this metric since a similar question was not contained in the March 2001 internal survey. This additional question was added to the December 2002 internal survey whereby 74% of the respondents agreed that the ROP is understandable and the procedures and output products are clear and written in plain English.

This metric met the criteria with a stable positive perception over time.

O-7 Public Perceives the ROP Maintains Safety

Definition: Annually survey external stakeholders through a *Federal Register* notice asking if the ROP adequately assures that plants are being safely operated and maintained.

Criteria: Expect stable or increasingly positive perception over time.

Analysis: Utility stakeholders believe the ROP maintains safety while a majority of non-utility stakeholders state it does not. The recent findings at Davis-Besse dominate the negative comments on this topic. While past surveys have had some negative comments on the ROP maintaining safety, this survey marks an increase in the proportion of negative comments.

This metric is not meeting its criteria due to a decreasing positive perception.

O-8 NRC Perceives the ROP Maintains Safety.

Definition: Annually survey internal stakeholders.

Criteria: Expect stable or increasingly positive perception over time.

Analysis: One internal survey question addressed this metric. However, no comparison could be made for this internal survey question addressing this metric since a similar question was not contained in the March 2001 internal survey. Among the December 2002 internal survey respondents, 76% agreed that compared to the previous process, the new ROP generally maintains safety.

This metric is indeterminate since the information could not be compared to previous internal survey results. However, internal stakeholders generally agreed that the ROP maintains safety.

O-9 Analysis of NRC's Responses to Significant Events

Definition: Review reports from incident investigation teams (IITs) and augmented inspection teams (AITs) to collect lessons learned regarding ROP programmatic deficiencies (e.g., did the baseline inspection program inspect this area, did the SDP accurately characterize resultant findings). IITs already have the provision to determine NRC program deficiencies. AITs will be reviewed by IIPB to identify any weaknesses.

Criteria: Expect no major programmatic voids.

Analysis: No IITs were conducted during the 2002 ROP cycle. Two AITs (Callaway and Davis-Besse) were conducted. IIPB reviewed the AIT reports and received a feedback form based on lessons learned from the Callaway AIT. Accordingly, IIPB revised IP 93800, "Augmented Inspection Team," and 71153, "Event Followup," regarding (1) AIT documentation requirements, (2) evaluating and documenting inspection findings provided they do not interfere with the AIT charter, and (3) risk metrics for events and degraded conditions.

In addition to the Davis-Besse AIT, the Davis-Besse Lessons Learned Task Force (DBLLTF) was formed to review the NRC's regulatory process associated with the issues at Davis-Besse. The DBLLTF's report, issued on September 30, 2002, contained more than fifty recommendations, many of which were associated with the ROP. Among the more significant ROP-related recommendations were: (1) enhance the barrier integrity performance indicators to more accurately measure unidentified leakage, (2) modify the inspection program to provide for better follow-up of longstanding issues, (3) develop specific guidance to inspect boric acid control programs and vessel head penetration nozzles, and (4) enhance the guidance for managing plants that are outside the ROP and under the NRC's IMC 0350 process.

Based on the programmatic deficiencies identified by the DBLLTF, the performance criteria for this metric was not met.

O-10 Analysis of Significant Events

Definition: Annually review all accident sequence precursor (ASP) events that have a risk significance of more than 10^{-6} to identify any ROP programmatic voids (i.e., did the baseline inspection program inspect this area, did the SDP accurately characterize resultant findings, etc).

Criteria: Expect no major programmatic voids.

Analysis: The Office of Research compared ASP results and SDP evaluations for several plant issues identified during the assessment period. No significant differences between the SDP findings and the ASP results were identified. During the period, several ASP reviews were initiated and the preliminary ASP for Oconee Units 1, 2, and 3 (high pressure injection and station auxiliary service water pump capability following postulated tornado) was completed.

The ASP program also evaluated selected operating events from the pool of licensee event reports, NRC inspection reports, and 10 CFR Part 21 notification letters. No potential accident sequence precursors (equivalent to or greater-than-green SDP findings) from this pool of operating experience information were identified during the assessment period.

The performance criteria for this metric was met.

O-11 Public Perceives the ROP to Be Effective, Efficient, Realistic

Definition: Annually survey external stakeholders through a *Federal Register* notice asking specific questions (based on NRC Strategic Plan) regarding whether the ROP is effective, efficient, and realistic.

Criteria: Expect stable or increasingly positive perception over time.

Analysis: In previous surveys, most if not all of the stakeholders stated that the ROP was effective, efficient, and realistic, or at the very least was a significant improvement over previous programs. From this survey, a distinct difference in perception is obvious between those utility stakeholders under the ROP and non-utility stakeholders evaluating the ROP. The majority of utility stakeholders have a positive response to this topic while the non-utilities believe the ROP has problems in this area. A common concern among most respondents is the efficiency and realism of the SDP.

This metric is not meeting its criteria due to an increasing negative perception.

O-12 NRC Perceives the ROP to Be Effective, Efficient, Realistic

Definition: Annually survey internal stakeholders asking specific questions (based on NRC Strategic Plan) regarding whether the ROP is effective, efficient, and realistic.

Criteria: Expect a stable or increasingly positive perception over time.

Analysis: Three internal survey questions addressed this metric. The percentage of NRC internal stakeholders agreeing that the ROP is effective (56% in 2002 vs. 57% in 2001) and efficient (70% in 2002 vs. 75% in 2001) has reduced slightly in both cases when compared to the previous survey, but remain relatively positive. However, no comparison could be made for this internal survey question addressing whether stakeholders perceived the ROP to be realistic since a similar question was not contained in the March 2001 internal survey. Among the December 2002 internal survey respondents, 65% agreed that compared to the previous process, the new ROP generally increases realism, and 74% agreed that the ROP provided a realistic approach to oversight.

The data supporting this metric indicates a slightly decreasing positive perception when compared to the previous survey. Although this metric did not explicitly satisfy the established criteria, internal stakeholders generally agreed that the ROP is efficient and realistic, and a majority agreed that the ROP was effective.

O-13 Resources are Commensurate with Performance

Definition: Correlate resources expended to Action Matrix column. Use RPS data to compare expended inspection resources to Action Matrix column by plant. Report high, low, and average.

Criteria: Expended resources should increase as licensee performance degrades (as noted by Action Matrix column). Establish baseline during first year of ROP.

Comments: This metric is no longer tracked and is being discontinued because it does not provide additional useful insights. Under the ROP, all plants receive the same level of baseline inspection. Inspections beyond the baseline are performed at plants with performance below established thresholds, as assessed through information gained from performance indicators and the results of baseline inspections. The ROP applies increasing inspection resources commensurate with declining plant performance. A correlation between increasing expended inspection resources for plants in successive reduced level of performance in the Action Matrix exists in the ROP by design and has been confirmed during the first two cycles of ROP implementation.

Analysis: Not applicable. This metric is no longer tracked and will be removed from the self-assessment program in the next revision to IMC 0307.

O-14 Public Perceives the ROP Enhances Public Confidence

Definition: Annually survey external stakeholders through a *Federal Register* notice asking if the ROP enhances public confidence.

Criteria: Expect stable or increasingly positive perception over time.

Analysis: Similar to previous surveys, the perception of public confidence in the ROP is a divided issue. Many respondents cite public participation in development and the consistent application of the ROP as major enhancements to public confidence. On the other hand, many stakeholders believe there is ambivalence within the general public towards the current system. Additionally, some stakeholders believe implementation and communication problems with the ROP are eroding public confidence.

This metric meets its criteria with a stable perception.

O-15 Opportunities for Public Participation in the Process

Definition: Annually survey external stakeholders through a *Federal Register* notice asking if there are sufficient opportunities for the public to participate in the process.

Criteria: Expect positive responses or an improving trend over time.

Analysis: As with the past surveys, a majority of respondents believe there is adequate opportunity for the public to participate in the ROP. Many stakeholders express concern that the public as a whole is not seizing those opportunities to provide input to the program. Additionally, some stakeholders believe barring the public from security based issues within the post 9/11 ROP is damaging its credibility and effectiveness on those issues.

This metric meets its criteria with mostly positive comments.

O-16 The Public Perceives the NRC to Be Responsive to its Inputs and Comments

Definition: Annually survey external stakeholders through a *Federal Register* notice asking if the NRC is responsive to the public's inputs and comments.

Criteria: Expect positive responses or an improving trend over time.

Analysis: The majority of respondents feel the NRC is responsive to inputs and comments, but others feel the NRC has no interest in outside input. This division is similar to previous surveys with most of the negative comments being either based on speed of resolution or a feeling that inputs are ignored or not expressly addressed.

This metric meets its criteria with a stable proportion of positive comments.

O-17 Public Perceives the ROP Was Implemented as Defined

Definition: Annually survey external stakeholders through a *Federal Register* notice asking if the ROP has been implemented as designed.

Criteria: Expect stable or increasingly positive perception over time.

Analysis: Respondents believe the ROP is being implemented as defined. A few stakeholders expressed concern that the actual documents are difficult to compile as a coherent reference. Specifically, numerous stakeholders agree that the practice of issuing preliminary non-green findings may not be in accordance with the program documents. Previous surveys indicated a similar agreement with only a couple of concerns with progress in revisions and aspects of the SDP implementation.

This metric meets its criteria with a stable positive perception.

O-18 Public Perceives the ROP Reduces Unnecessary Regulatory Burden

Definition: Annually survey external stakeholders through a *Federal Register* notice asking if the ROP reduces unnecessary regulatory burden.

Criteria: Expect stable or increasingly positive perception over time.

Analysis: As with the previous surveys, stakeholders believe the ROP does reduce unnecessary regulatory burden. In addition to that feeling, a few believe the program may be going too far and reducing what is actually necessary regulatory burden. Some utility respondents feel the SDP process can be further refined to reduce the time and resource expenditures in its screening process.

This metric meets its criteria with a stable public perception.

O-19 Public Perceives the ROP Does Not Result in Unintended Consequences

Definition: Annually survey external stakeholders through a *Federal Register* notice asking if the ROP results in unintended consequences.

Criteria: Expect stable or increasingly positive perception over time.

Analysis: The majority of stakeholders responding indicate that they believe the ROP resulted in some unintended consequences. Examples include the recent vessel head issue at Davis-Besse and the assigning and subsequent changing of the preliminary color of SDP results creating an undue concern. Currently, the overall perception of unintended consequences is similar to previous years. While the actual consequences have changed, the number and gravity of their impact has remained constant.

Although the results are similar to previous years, these results do not meet the staff's expectations. Therefore, this metric has not been met.