Official Transcript of Proceedings ACRST-3275

NUCLEAR REGULATORY COMMISSION

ORIGINAL

Title: Advisory Committee on Reactor Safeguards Reliability and Probabilistic Risk Assessment & Plant Operations Subcommittees

Docket Number: (not applicable)

PROCESS USING ADAMS
TEMPLATE: ACRS/ACNW-005

Location: Rockville, Maryland

Date: Wednesday, April 14, 2004

Work Order No.: NRC-1413

Pages 1-121

NEAL R. GROSS AND CO., INC.
Court Reporters and Transcribers
1323 Rhode Island Avenue, N.W.
Washington, D.C. 20005
(202) 234-4433

TR04

UNITED STATES OF AMERICA

NUCLEAR REGULATORY COMMISSION

+ + + + +

ADVISORY COMMITTEE ON REACTOR SAFEGUARDS

MEETING OF THE

RELIABILITY AND PROBABILISTIC RISK ASSESSMENT

AND PLANT OPERATIONS SUBCOMMITTEES

MITIGATING SYSTEMS PERFORMANCE INDEX

+ + + + +

WEDNESDAY, APRIL 14 , 2004

+ + + + +

ROCKVILLE, MARYLAND

The Subcommittees met at the Nuclear Regulatory Commission, Two White Flint North, Room T-2B3, 11545 Rockville Pike, at 8:30 a.m., John D. Sieber, Chairman, presiding.

COMMITTEE MEMBERS:

JOHN D. SIEBER	Chairman
MARIO V. BONACA	Member
F. PETER FORD	Member
THOMAS S. KRESS	Member
GRAHAM M. LEITCH	Member
STEPHEN L. ROSEN	Member
WILLIAM J. SHACK	Member
MAGGALEAN W. WESTON	Staff Engineer

NRC STAFF PRESENT:

DAVID ALLSOPP

JAMES ANDERSON

PATRICK BARANOWSKY

BRUCE BOGER

BENNETT BRAG

DON DUBE

ANNE-MARIE GRADY

NAOTO ICHII

AUDREY KLETT

PATRICK O'REILLY

SELIM SANCAKTAR

JOHN THOMPSON

C-O-N-T-E-N-T-S

P-R-O-C-E-E-D-I-N-G-S

8:28 a.m

DR. SIEBER: The meeting will now come to order.

Good morning. This is a meeting of the ACRS Subcommittees on Reliability and PRA and on plant operations. I'm Jack Sieber, Chairman of the Plant Operations Subcommittee.

George Apostolakis -- and I don't see him here -- is the Chairman of the Reliability and PRA Subcommittee.

Other ACRS members in attendance are Mario Bonaca -- and he's here but not at the table, Peter Ford, Tom Kress, Graham Leitch, Steve Rosen and Bill Shack is, I think, is supposed to be, too. Okay?

The purpose of the meeting is to discuss the technical results of the Mitigating Systems Performance Index Pilot Program. Maggalean Weston is Cognizant ACRS Staff Engineer for this meeting.

The rules for participation in today's meeting have been announced as part of the notice of this meeting published in the Federal Register on March 24, 2004.

A transcript of the meeting is being kept and will be made available as stated in the Federal

Register notice. It is requested that speakers use one of the microphones available, identify themselves, and speak with sufficient clarity and volume so that they can be readily heard.

We have received no written comments from members of the public regarding today's meeting.

I think now we'll proceed with the meeting. Pat Baranowsky of the Office of Research will begin.

MR. BARANOWSKY: Thank you, Mr. Sieber, and good morning members of the ACRS Subcommittee. Thank you for this opportunity to come here and talk to you about the technical evaluation that we had performed with regarding to the Mitigating System Performance Index development over the past two years.

If you can go to the purpose slide. First we're going to spend the bulk of this meeting talking about our evaluation. This is the third meeting that we've had with this subcommittee on this topic. And now we're coming toward the end of this project.

We'll have a brief presentation by NRR on the status of the MSPI and there has been a request at least for one public member to make a short presentation also on the MSPI by NEI. Some time has been allotted for that.

Another underlying purpose of this meeting is that we would like to ultimately get a letter at the conclusion of this project. We actually talked about this back in I think it was May of 2002. And we're actually somewhat on schedule believe it or not, looking for that letter a little bit later this year.

Let me start off by giving our overall conclusions as we see it on this work. Recognize, of course, that the work to date is presented in a draft report that was provided to you and other members of the staff in February of this year. It was also released to the public about a week or two ago through a Federal Register notice. And so it's still not in its final form but it's getting very close.

As a result, we think that the Mitigating System Performance Index is a highly capable performance indicator that can differentiate risk-significant changes in performance and address problems associated with the current performance indicators.

The development activities --

DR. LEITCH: Pat, are you going to refresh us on just what are the problems with the current performance indicators?

MR. BARANOWSKY: On what the current ones are?

DR. LEITCH: Yes, I'd just like --

MR. BARANOWSKY: Okay.

DR. LEITCH: -- to be a little more clear as to what --

MR. BARANOWSKY: Okay.

DR. LEITCH: -- we're trying to solve here.

MR. BARANOWSKY: If it's not in there, we can -- we can --

MR. DUBE: It's in a --

MR. BARANOWSKY: Oh, the problems, yes, we can have that. That's in there.

DR. LEITCH: Okay.

MR. BARANOWSKY: I'm sorry. I thought you wanted to know what the current indicators were.

DR. LEITCH: No, what the issues are.

MR. BARANOWSKY: No, we have those.

DR. LEITCH: Okay, thanks.

MR. BARANOWSKY: I'm just getting to the conclusions to you'll know what we're going to try and present.

We've had an extensive developing testing program and for the most part completed our evaluation, looking at its validity and verifying its capability. We think we understand its performance characteristics, its strengths, and its limitations very well.

To us it appears to provide the best overall measure of system performance while minimizing false positive and false negative performance indications. And this is especially true for identifying changes in performance.

Also please note that the formulation is flexible and adaptable and, in fact, it's been modified substantially from its original formulation almost two years ago. And, as such, we've been able to address and can continue to address emerging issues and concerns regarding validity and appropriateness of the outcomes using this indicator.

Next -- so the RES presentation here, which will be followed by NRR and some public comments is as follows. We'll go through the background, an overview of what the MSPI is, the status of the pilot program, and scope of our verification activities.

We'll discuss the research results of the pilot program, some key technical issues that have received significant activity on our part and other members of the NRC. In particular, we'll talk about the validity and robustness of the MSPI and give you our overall assessment and conclusions in a little bit more detail.

Next one -- the MSPI evolved from a feasibility study of Risk-Based Performance Indicators that was done by the Office of Research and documented in NUREG-1753. I think we started that work about four years ago. And we actually had several meetings with the ACRS subcommittees and the full committee on that work.

As a result, when some problems were identified with the current set of performance indicators for mitigating systems, NRR came to us and asked if we could adapt that work to solve those problems. And the third bullet here pretty much identifies what the issues were that were identified by NRR.

The use of fault exposure time is a surrogate for unreliability. The definitions of unavailability were inconsistent with the maintenance rule and actually in some cases, inconsistent with PRA usage.

There was cascading of cooling water support system failures on to front line systems and concern about how that would impact the way the performance indicators were used in the action matrix.

In the thresholds, and the indicators themselves were not plant specific but generic, one size fits all, and there was a significant concern about plant-specific differences.

DR. LEITCH: Could you say a little more about that third bullet? That cascading of cooling water support system failures?

MR. BARANOWSKY: Yes.

DR. LEITCH: What I'm picturing is say you have a -- an RHR pump that needs cooling water to the bearings.

MR. BARANOWSKY: Right. The current set of performance indicators has about four or five front line system performance indicators and, of course, each hit on one of those systems produces a color indication which then goes into the action matrix if you achieve certain levels.

So if a cooling water system is found to have a fault that effects two or more of those front line systems, then each system is credited as having a hit and, therefore, you might get two or three performance indication hits when there is actually one system that's the problem. And so we're trying to correct that.

DR. SIEBER: On the other hand from a risk standpoint, if you lose cooling water, you lose a lot of pumps, you lose your diesels --

MR. BARANOWSKY: Right.

DR. SIEBER: -- I would say it's risk significant.

MR. BARANOWSKY: Yes.

DR. SIEBER: And I'm not sure that taking a bunch of hits is a wrong thing.

MR. BARANOWSKY: We're not saying it's not risk significant. But remember you can have a single PI hit that goes anywhere from green all the way up --

DR. SIEBER: Right.

MR. BARANOWSKY: -- to very significant red. The numbers of PI hits is meant to indicate how many systems and components are effected so you can understand the breadth of the issue. The color up to red is meant to give you the significance of the individual findings.

And so we want to not confuse that philosophy here with the performance indicators. And I think that was pretty universally agreed upon that we should go in that direction.

And, by the way, the current formulation allows us to look at the significance of cooling water

-- when I say currently I mean Mitigating System Performance Index, allows us to look at the significance of those cooling water systems with respect to their risk-significant safety function so we don't lose that aspect.

Okay, so first we went through a -- what I would call a modification and development phase in which we took the formulations from NUREG-1753 and came up with the basic or fundamental formula for the Mitigating System Performance Index. And then a 12-month Pilot Program was initiated in September 2002 to test out and evaluate the Mitigating System Performance Index.

We briefed the ACRS; the last time on this was July 2003. We covered some issues that were raised in a May 2002 subcommittee meeting. And, as a result, I believe we answered all the questions that were raised and no significant new ones were raised although we said we would come here at this time and let you know the results of the Pilot Program. And so here we are.

DR. LEITCH: Now a 12-month Pilot Program implies that you factored actual operating experience into this pilot to see how this indicator would react. Could you not also have just assumed certain failures to see how the indicator would react? I don't understand -- I guess I don't understand that about the pilot.

MR. BARANOWSKY: In fact, that's a great point because we did do that. What we did was both the Pilot Program where we used the actual operating experience and found out what it was like to collect the information and handle it and --

DR. LEITCH: Okay.

MR. BARANOWSKY: -- make calculations.

DR. LEITCH: All right.

MR. BARANOWSKY: And then we did numerous simulations in which we simulated --

DR. LEITCH: Certain failures?

MR. BARANOWSKY: -- the operating experience so that we could really understand the implications of different changes to the MSPI formulation. And that, I think, is one of the key parts to our ability to develop and understand the performance indicator.

MR. DUBE: Yes, I might go into that a little later. But we did do Latin Hypercube simulation of failures. We assumed the distribution of failure rates for the various components. And then just simulated -- like a Monte Carlo simulation.

DR. LEITCH: Was any of that historically -based? Maybe we'll talk about that a little bit later but in other words, did you take a look at hey, here's a pretty serious event that happened at Plant X back in 19-something or other and --

MR. DUBE: No, no --

DR. LEITCH: -- factor that into the program and see if it gave you the right color?

MR. DUBE: Well that aspect we did. But the simulation didn't do that.

DR. LEITCH: Okay.

MR. BARANOWSKY: There were some special cases where we looked at specific incidents especially to see whether or not the indicator would have been the appropriate tool to take a look at that particular condition.

Very shortly I'm going to get to some of the limitations --

DR. LEITCH: Okay.

MR. BARANOWSKY: -- and that's important to understand those, too.

DR. ROSEN: The main -- the other half of that question was you explained that it was -- it also gave you an opportunity to see how difficult it was to collect the data. Are we going to hear more about that?

MR. BARANOWSKY: We weren't planning on going into that in too much detail but we can talk a little bit about it.

DR. ROSEN: Maybe the industry is going to talk about that?

MR. BARANOWSKY: Well, I think the industry can tell you about how difficult it was on their part.

DR. ROSEN: Yes.

MR. BARANOWSKY: Because from our part it wasn't very difficult.

DR. ROSEN: I want to hear what -- well both sides of that story.

MR. BARANOWSKY: Okay, all right.

So as I mentioned, we did formulate the indicator that eliminates the specific problems that were identified. It addresses those.

It accounts for unavailability and unreliability in a system weighted to its relative risk importance, uses a plant model to derive risk importance weightings. In other words, it's plant specific.

It identifies changes in performance while limiting false positive and false negative indications which is an issue that I hope we'll get an opportunity to go through a little bit more when Don starts talking.

And lastly, it's quite consistent with PRA maintenance methods and the maintenance rule data collection. And as best we understand it, having interacted with not only the group that was part of the pilot but also with INPO, we -- a system is being set up called the consolidated data entry which is encompassing the currently existing EPIX, that's the Equipment Performance Information Exchange System which could capture the data necessary to generate the MSPI.

Next -- okay, as I mentioned, the indicator monitors basically changes in performance that are related to changes in core damage frequency. We call it an index because it's really only a partial indication of changes in core damage frequency. It doesn't include everything within that system that could result in a change in core damage frequency and about two slides away, I'll tell you what we don't cover.

There are two elements in a -- because of the formulation can be combined very simply and linearly. That was a significant problem with the

NUREG-1753 work where we had come up with two indicators for each system; one for unreliability and one for unavailability.

We couldn't combine them together because at that time, we didn't put them into a core damage frequency common denominator approach, if you will, and, therefore, the significance of each of those wasn't properly weighted when we first formulated it although we knew we could probably do it.

And thus we have an indicator that has the two parts, as I mentioned, the unavailability and the unreliability index, which is related to the change in core damage frequency associated with change in unreliability and unavailability.

And what we were able to do was to use basically the Fussell-Vesely importance measures to linearize the whole process and make it fairly simple once one has a PRA to work with.

DR. SHACK: Pat?

MR. BARANOWSKY: Yes?

DR. SHACK: One of the interesting things was the variability you got in the Fussell-Vesely for components, which was of interest to me, of course, in a 5069 kind of sense. And I notice in Appendix B, this is attributed to the fact that your models did or did not have initiating event fault trees.

Can I draw a general conclusion that if I'm looking at PRAs that don't have initiating event fault trees, I'm computing suspect Fussell-Vesely numbers --

MR. DUBE: They could be off by --

MR. BARANOWSKY: Could be.

MR. DUBE: -- a significant amount.

DR. SHACK: Yes, I mean these were big changes.

MR. DUBE: Yes.

MR. BARANOWSKY: Yes.

MR. DUBE: If a particular special initiator, let's say loss of service water is a dominant sequence of the plant and it involves the failure of pumps and valves and components thereof, and one PRA model has an explicit fault tree that's for that initiating event that's linked with the rest of the model and another one uses I'll say a single-parameter frequency, you could have significant differences in the importance measure. Yes, order of magnitude we saw, in some cases even more than an order of magnitude. That was an eye-opener when we --

DR. SHACK: Yes, I thought that was pretty impressive.

MR. BARANOWSKY: Yes, you know I've been around here for over 30 years. And I think -- maybe I don't have a gray beard but I qualify as a gray beard. And I learned a few new tricks on this project about how sophisticated one needs to be with PRA to capture results that present persistent outcomes in your conclusions.

And it takes a little bit more sophistication than just getting the top number correct so to speak.

MR. DUBE: Right, yes. The interesting thing is that the core damage frequencies can match between the two models but the importance measures can be very different.

DR. SHACK: Very different, yes.

MR. DUBE: And that's why one of the recommendations was to be -- put all plants on an equal footing, one needs to address this issue of support system initiators. And that it is one of the recommendations to do that.

DR. ROSEN: I guess requiring modeling of the support systems but that's not an impossible task.

MR. DUBE: No, it isn't.

DR. ROSEN: It's really, in fact, fairly straightforward.

MR. DUBE: Correct.

MR. BARANOWSKY: In fact, I think we offer a --

DR. ROSEN: An alternative --

MR. BARANOWSKY: -- a simplification based on having studied this that would allow plants without a fault tree to come up with appropriate results with their support system initiators.

DR. ROSEN: I put that in the category of less than a full scope PRA. If somebody's just using a plug and jug number rather than modeling the support systems, it's just another one of those examples that the PRA folks didn't finish the work.

MR. BARANOWSKY: Well, sometimes we're guilty of doing that.

DR. SHACK: Well, I just want to make sure it's captured when we do a 5069.

MR. BARANOWSKY: Yes.

DR. SHACK: That was definitely my concern.

DR. ROSEN: Keeping in mind when you think about whether or not an applicant or a licensee has a full scale PRA if we're trying to judge that.

MR. BARANOWSKY: That's one of the -- if not the highest item on our SPAR upgrade list, by the way, that came out of not just this project but having gone through now and looked at every single PRA, every single PRA, and comparing them to SPAR. And so we think we need to improve in that area in order to get the PRA results correct.

Okay, I'm not going to go over the formulation any further here because it's developed in detail and we have also presented this at the prior meeting. But I think it's kind of elegant and simple and yet it does a lot. And I'll just leave it at that.

DR. LEITCH: One question that I had -- just back to the previous slide there if you could a second. I'm always concerned that when we have performance indicators, we begin -- or as an industry to manage those indicators. And sometimes that can yield some unintended consequences. Is there something in this formula that would cause the utility to want to drive the UAI as low as possible?

MR. BARANOWSKY: Sure --

DR. LEITCH: I guess one of the things that always concerns me about this --

MR. BARANOWSKY: -- I think they would want to drive the unreliability and unavailability low. And if they do that, I'm all for it.

This is one of those cases where we need -- I agree with you on that. Performance indicators can be set up so that people implementing them worry about the performance indicator -- and it will happen.

In this case, the performance indicator is so closely linked to plant risk, that it's a good thing. It's a good thing to have low unreliability.

DR. ROSEN: As long as you balance it.

MR. DUBE: Absolutely.

DR. ROSEN: But my concern --

MR. BARANOWSKY: Well, that's why you see both in there.

DR. ROSEN: My concern is if one tries --

MR. BARANOWSKY: It's a great -- from that point of view.

DR. ROSEN: -- if one tries to drive the unavailability to zero, for example, it can adversely effect the unreliability because you're not taking the time required to do the proper preventative maintenance and those types of things. So --

MR. BARANOWSKY: Agreed.

DR. ROSEN: -- there is kind of a balance between those two terms.

MR. DUBE: I agree. And that's why in the current situation, you only have the first term in this equation, the UAI part.

DR. LEITCH: Right.

MR. DUBE: And you can see -- and there have been examples where for whatever reason, a licensee is right on the borderline and has managed unavailability in order not to cross for better or for worse.

DR. LEITCH: Yes.

MR. DUBE: In this case, you know, reliability theory says you want to optimize your preventative maintenance to give you the best combination of unavailability and unreliability.

Too much maintenance and the UAI term goes up, you know, the URI term may go down to zero but that's not optimum. Too little maintenance, UAI goes to zero and URI can shoot up.

DR. LEITCH: Right.

MR. DUBE: And the best world is the right combination of unavailability and unreliability. And better yet, my belief is that the MSPI weights unavailability and unreliability based on your risk importance.

DR. LEITCH: Yes, yes.

MR. BARANOWSKY: Which makes it consistent with the maintenance rule which basically says balance these things out.

MR. DUBE: Yes.

MR. BARANOWSKY: Okay. The systems that are currently monitored or capable of being monitored by the MSPI are indicated here. We all know what these systems are. One could expand, if one wanted, very easily due to the formulation, add different systems, different scopes and so forth.

It could even expand this to initiating events if you wanted to because it has such a general applicability. But I don't plan on going into the details for the scope of the systems here right now.

DR. ROSEN: But if you were to do that, give us a feel for how many -- how much more percent of the CDF you would get? Can you do that? Or do you think this is half of -- this covers half of the CDF? Or 75 percent? Or 90?

MR. BARANOWSKY: Well, it covers a large chunk of the system components that are involved in the CDF. It indirectly includes initiating events in that, of course, the current CDF is based on what the current initiating event of that frequency is but it doesn't account for changes that might be occurring in the current initiating events.

So that portion of risk that might be

changing as the result of changes in initiators aren't, obviously, captured here. That's another milestone.

There are aspects of the MSPI and they're actually on the next chart that we're unable to account for in the current formulation -- or at least we haven't really tried to hard to do this. And why don't we just flip to that because it is important to know what the limitations are.

DR. ROSEN: Are you going to try to answer my question?

MR. BARANOWSKY: I'm going to try to answer it when I get to the end if you'll -- if I get to this point.

DR. ROSEN: All right.

MR. BARANOWSKY: Multiple concurrent failures of components, including common cause failure, are not currently included in the formulation although the importance of common cause failure on the Fussell-Vesely and hence the total formula is included.

So in other words, changes in performance that are due to a greater susceptibility or the actual occurrence of multiple failures, including common cause, are not within the scope of the current MSPI.

Also, conditions that are latent and not discovered by routine surveillance and, therefore, can be in existence for several surveillance intervals and may require a design review or some special test to detect, they're not included.

And lastly, failures of passive components are also not included. So what we're looking at is the key contributors to risk from an active component point of view and the aspects of those components that go beyond the capability of the MSPI would be covered by a significance determination process activity.

Now I can't say how much CDF is accounted for but for -- I would say that the CDF associated with the MSPI is not the largest chunk of core damage frequency that would be found in the PRAs although the reliability of equipment that's within the scope of the MSPI can be found in some of the dominant sequences.

Now that sounds like a little bit of double talk. But in essence, remember what we're talking about here is single failures of components that are detected during normal surveillance and what the implications are of those failures on performance and risk.

Generally that's not the largest

contributor to core damage frequency. Larger contributors are associated with common cause failure and some other factors that are not easily monitored through system reliability monitoring.

DR. ROSEN: Let me -- let me restructure my question to you --

MR. BARANOWSKY: Okay.

DR. ROSEN: -- to get at really what I was asking because, sir, I think you're right about what you just said.

If you were to formulate a list of systems to cover under MSPI and be inclusive, would there be additional systems on this list? And if so, how important would those additions be to the result?

MR. DUBE: I mean -- if I can answer -- you've got basically high pressure safety injection here and, to a certain extent, residual heat removal where it's shared with low pressure safety injection. You've got that covered.

You've got emergency feedwater and reactor core isolation cooling. And if we don't have -- we didn't have an isolation condenser plant in here but that would be included if -- in here as well. And then you've got your major support system service water and component cooling water and emergency AC power.

So in answer to your question, the only thing I can think of would be something like a DC power, you know 120-volt AC and maybe for some plants where instrument error is important, that might be one. But I would say we've got the bulk of, you know, the important systems already here.

DR. ROSEN: Okay, well that's what I wanted to hear.

MR. BARANOWSKY: We actually -- NUREG-1753 showed that the coverage that we have is very large. And our philosophy is that -- remember this is a sample of performance, the theory being if we sample enough things in the most important areas and they're not going well, that's indicative of other things that are not easily sampled.

It's not easy to sample common cause failure things. I don't think anyone knows how to do that. But we do know there is somewhat of a correlation between common cause failure and independent failures.

You have very, very few independent failures. Your common cause failure contribution to your risk is generally very low. So there is a relationship but it's not a hard and fast one.

In theory, if we have problems with these systems due to these single failures, they're sort of a gateway, if you will, into what else might be going on.

DR. FORD: Do I understand it from the previous graph that, for instance, aging effects on passive components are not covered in this overall scheme of events? And if they are not, will they be in the future?

MR. BARANOWSKY: They're not covered here. There is some thought being put into developing performance indicators that would be related to that issue. And I'm not sure whether we will or will not go forward with that activity. There is some work that is scheduled for the next fiscal year on that.

DR. FORD: Okay, good.

MR. DUBE: If I might add, if a passive component, let's say a heat exchanger or some piping section caused a train in one of the monitored systems to be unavailable, that would get captured in the MSPI because we capture train unavailability.

But if it was a catastrophic failure and resulted in a leak or in an initiating event, it would not. But it would default to the significance determination process. It would be covered there.

DR. FORD: Would that leak and therefore of the availability of that passive component would be a reactored step? You'd wait for the leak to occur before you started to fit it into your analysis. Is that correct?

MR. BARANOWSKY: Yes.

MR. DUBE: Yes.

DR. FORD: Okay.

MR. DUBE: In looking at its impact on unavailability.

MR. BARANOWSKY: That's one of the reasons why I say I'm not sure where we're going to go with this because we don't want to just track pipe breaks --

DR. FORD: No.

MR. BARANOWSKY: -- that's just not really a good level of tracking just like tracking common cause failures and waiting until you have a dozen of those isn't a good idea either.

So we're trying to look at whether or not there is some sort of condition monitoring aspect of performance that might be used instead. And I don't know where we're going to go with that.

DR. FORD: Okay, but at least it's in your thought pattern.

MR. BARANOWSKY: It's in -- sort of in the thought process. If we think it has any practicality, we'll go further with it. It may need some more research on monitoring of equipment.

DR. SHACK: But that's one of the problems I always sort of have with one of these integrated approaches in the first place is that, in fact, you smear the performance out because you're sort of giving them credit for all the options they have of mitigating a problem.

Even if they have a problem, it doesn't show up as serious here because you're -- you know, you're crediting the other mitigation strategies that are sort of inherent in the plant. And while that's true if I was looking for a, you know, a true risk impact of this, but in a performance measure, I'm measuring more than risk impact, I think.

You know I'm trying to look for a precursor. And it seems to me as I keep integrating my performance indicator, I'm losing something of the performance indicator and I'm getting much more of a safety indicator, which is of interest in itself but I lose -- you know, I gain and I lose by combining these systems together the way that you have.

MR. BARANOWSKY: Well, one of the

philosophies -- and I was going to bring this up a little bit later -- is that licensees are supposed to take the maintenance rule, for instance, and do the early screening of these very issues that you're talking about. And when the performance indications get to a certain stage, then the NRC steps in.

And so what we're trying to do is identify that stage based on its risk significance.

DR. SHACK: Yes, but see I look at this -- when this performance indicator is bad, things are very bad. I mean --

MR. BARANOWSKY: I think I would take exception to that.

DR. SHACK: If things have gotten to that stage, then they're very bad. When this performance indicator is good, I'm not so sure that things are good.

MR. BARANOWSKY: I guess I would disagree with you.

DR. SHACK: Okay.

MR. BARANOWSKY: I think that we have to recognize that we're looking at changes in performance that result in changes in core damage frequency.

DR. SHACK: You're looking at changes in risk. You're not looking --

MR. BARANOWSKY: Changes in core damage frequency on the order of 10 to the minus 6 when the total core damage frequency is closer to 10 to the minus 4 -- we're down two orders of magnitude from what might be the total baseline.

And so before I take and claim that things are very, very bad, I would want to look at the total risk perspective. We're actually attempting to work around the resolution in risk analysis here.

DR. ROSEN: I think this is the -- not the electron microscope for performance. The electron microscope, the thing that shows you the real fine structure is the maintenance rule because it's -- the licensee, if he gets more than a certain number of failures has to set up a program, put it in A1 or A2, I forget which one it is, and create a program to correct those problems on the individual component.

And so that -- and the NRC can see any time -- the resident can go look at what's on the list anytime. So I would rely on that for the fine structure rather than this program. This program is more step back and look at the forest rather than the individual trees is the way I see it.

MR. BARANOWSKY: Well, it's a pretty fine level of resolution though when you look at the whole picture. That's why we actually, as an Agency, broke it up into small pieces instead of just saying, "Let's just look at core damage frequency." We're going down and taking all these small pieces of core damage frequency and we're not taking the total core damage frequency.

If we were, we would be looking for changes of 1 to .1 percent in total core damage frequency. I have a nice little picture to show you what that really means later if you want to see it.

Okay any more questions on that one?

(No response.)

MR. BARANOWSKY: I think this is where I turn it over to you, Don, is it?

MR. DUBE: Either way.

We kind of touched upon many of this. The 12-month pilot was completed in September of last year. We did have a preliminary draft report to what has been distributed here.

In the interim from September through pretty much January, we researched and our contractors did some additional analyses. And I'll be touching upon those particularly with regard to PRA adequacy and some comparison of results.

We continued to hold public meetings.

Internally, NRC meetings were held and position papers formulated. We issued the draft report on the verification that you have a copy of.

And then last month, the Division of Inspection Program Management terminated the development and implementation of MSPI and they'll discuss that in a little while.

The independent verification -- this was the original scope of what we were intended to do. And it was pretty comprehensive quite frankly. We verified all the baseline data, reviewed all the unavailabilities, we did note some inconsistencies and they're highlighted in the report. But in general, they were pretty reasonable.

We revised the industry failure rates to represent most current performance, which is a little bit better than the period '95 to '97 but within statistical uncertainties is represented of it whereas the failure rates we originally had for this program was 10 and 20 years old. So this was an important improvement that we made along the way.

We verified all the performance data, all the unavailabilities. We compared all the reliability data of all the pilot plants for all the components to EPIX and in some cases to the reliability and availability database. We did not some errors and those were corrected as the program progressed.

The Fussell-Vesely importance are input to the process. We compared those to the SPAR models and there's a whole appendix in the report on that. But we did find substantial differences, especially in many of the support systems. And as a result, we had an unexpected and unanticipated SPAR enhancement effort where we improved 11 SPAR models down to the level of component risk importances.

And then we analyzed those differences between the licensee PRA model and then the SPAR model. And then we did sensitivity studies based on that which I'll touch upon in a few minutes.

We verified the spreadsheet -- that it was doing the calculations correctly. We compared the MSPI results using SPAR and the licensee's PRA so we had one-for-one comparison there.

And then we analyzed the differences. We performed sensitivity studies. And then as I mentioned, we analyzed the results for all the component failures in the pilot, which was some 77 failures.

So we went through 77 failures for the systems within the scope of the pilot and each one we said what did the MSPI result? And in a lot of cases, it was a cumulative effect so you have to look a little bit before and a little bit after, compared to what the equivalent SSU was for that quarter and if there was an SDP, what the SDP indication was.

And there wasn't always one. For example, out of the 77 failure, I believe there may have been 18 or 20 or something like that SDPs. So it was a very comprehensive, independent verification effort.

Now in terms of the research results, we were able to find very good agreement between the plant models and the SPAR resolution models. Now these are the SPAR models after we were -- we were -- made the adjustments and refinements.

And later on, I'll talk to you how we did sensitivity studies and backed off on those SPAR models to look at what would the impact be because of the differences. So it was pretty comprehensive.

But we were able to -- it's more than just a fine tuning and a benchmarking. It was understanding what the differences were between plant PRA model and SPAR models. Why were there differences in dominant sequences and cut sets and importance measures.

And what we found, as Mario -- not Mario -

- as Pat mentioned earlier was that we were able -- you know, the SPAR models were pretty accurate to begin with. We could predict core damage frequency to within factors or two or three. And we did pretty much have the dominant sequences and even the dominant cut sets.

But at the importance measure level, we found significant differences. And we couldn't stop there because those importance measures are what is input into the MSPI.

So we evaluated the differences in the model. For the 11 models, we found only three plant-specific model differences that could potentially have a large impact on the results. And I'll touch upon that. There were a number of others that had medium impact and a number that had small impact.

We found that the significant differences in major model inputs were such things as system success criteria or initiating event frequencies for major initiating event frequencies for support systems. They were the primary source of significant quantitative differences whereas when we looked at factors of two or three differences in basic event probabilities, they were generally almost always low impact on results.

So the licensee model said the probability of failure of a component was 10 to the minus 2 and the SPAR model might be two or three times greater. When you run that through the PRA models and it aggregates it and then run those importance measures that are generated into the MSPI, it made virtually no difference at the basic event probability level, which was an interesting result.

I discussed how we compared the MSPI, SDP, and SSU results for all 77 component failures. And I'll discuss the more important ones. There was some agreement and there was some disagreement. All our explainable but we do realize that SDP and MSPI do have fundamentally different purposes.

But it was a task that we were asked to do and so, you know, we did the best that we could in that with recognizing those differences.

The sensitivity studies were done to address PRA adequacy. In other words, let's assume that the licensee has a PRA model and there is a SPAR model. And there are difference in models. They might be differences on basic event probabilities but also it could be success criteria could also be, to some extent, fault tree and event tree structure.

We identified the major differences

between these for all 11 models and then we grouped them into somewhere between three and seven categories where it made sense. But generally we used seven categories.

So all the differences that would reflect themselves in emergency AC power, we grouped all those changes, all those differences together. And all the differences between the two models that effected aux feedwater system, we grouped those together. And those that effected, let's say PORV success criterion for feed and bleed were grouped together.

And then we created change sets. When I say we, it was primarily Idaho National Lab -- generated change sets and ran the SPAR model. We ran two at a time to see how these groups of differences -- - now let's say aux feedwater, there were four or five differences, how those would effect the PRA results.

And then -- so we generated new PRA results, including revised Birnbaum end points measures and Fussell-Veselys. And then we took those new Birnbaums and fed them back into the MSPI algorithm to generate new MSPI results.

So in short, we looked at all the major differences between the SPAR model and the licensee's PRA, grouped the differences in a logical fashion, reran the PRA -- let's see 20 -- somewhere well over 100 times, generated new Birnbaums, put those into the MSPI one set at a time and generated new MSPI results to look at how a little input difference here might manifest itself in a difference in MSPI results.

And the quantitative and qualitative changes in the MSPI provides a measure of the sensitivity of the results to model differences.

And the next slide summarizes the sensitivity studies. And we grouped them into three categories. And one might argue it's a little bit arbitrary. But we defined large as the difference is greater than 5E to the minus 7, recognizing that the lowest threshold at green light is 10 to the minus 6. So it's likely to effect to color performance indication given some failures in the system.

In order to do this, by the way, we assume all components have one failure beyond the baseline. That's a little parenthesis at the bottom. So it is a little bit conservative to begin with because we're saying when we did the sensitivity study let's assume that every component in that system had at least -- had one failure more than the baseline. So the 5E to the minus 7 really is kind of a conservative value, quite frankly.

The medium differences were between 10 to the minus 7 and 5E to the minus 7 and had the potential to effect the color. But you'd have to have a significant number of failures in the system to do that.

So finally there was the low, which was less than 10 to the minus 7' and very unlikely to effect or skew our results. And we did this so we could have a logical and consistent way of looking at the hundreds of differences between the licensee's PRA and the SPAR models and make some sense out of them and say what was important and what was not.

And the table below summarizes all of the differences and grouped into large, medium, and small. And the Braidwood PORV Success Criterion has to do with the fact that the SPAR model assumes feed and bleed that two PORVs are needed whereas the licensee's PRA, plant PRA, uses one.

It's not a judgement necessarily whether the licensee's PRA is correct or not because there are some indications that one PORV may be adequate for successful feed and bleed but it is a measure of the sensitivity of how differences in the models reflect themselves and could impact the MSPI.

Millstone 2 has to do with a number of

issues. They, in the last year or two, couple years, they changed the LOCA categories and thereby changed the frequencies. And we couldn't get them to match the more generic, if you will, combustion engineering 2,700 megawatt thermal kinds of LOCA categories and frequencies. And so there were differences there. And so they did manifest themselves in the large effect.

And the third one was Salem. It had to do primarily with the service water system initiating vent frequency where the Salem initiating event frequency is about 30 times lower than what is used in the SPAR model. And here my personal belief is that the licensee's frequency is lower than what one would generally determine to be a nominal value.

And that was it.

In the medium level, I'm not going to go through them in detail. But there were a number of issues. And then all the other -- all the other differences, literally 100 -- well over 100 differences had smaller or no effect.

DR. ROSEN: Now were there plants in the pilot beyond the ones that are listed on this slide? In other words, called out specifically?

MR. DUBE: Some that I didn't show up

here?

DR. ROSEN: Yes.

MR. DUBE: I think -- let's see --

DR. ROSEN: I mean with large or medium --

MR. DUBE: -- south Texas doesn't show up here, San Onofre doesn't show up here, Surry-1 and 2 don't show up here. So the answer is yes.

DR. ROSEN: And that is mainly, I think -- I'm asking if this is true, is that true because those plants have models that are very close to SPAR? Or SPAR is very close to their model?

MR. DUBE: It's a combination of that and a combination of the MSPI results are not sensitive to whatever differences there are.

MR. BARANOWSKY: They're close enough basically?

MR. DUBE: Yes, and to answer your question, they are pretty close.

DR. ROSEN: Okay, because this chart is all about how different SPAR is from the model at the plant, isn't it?

MR. BARANOWSKY: Yes.

DR. ROSEN: Okay. So I'm concluding that plants that were in the pilot that don't show up here have models that are pretty close to SPAR --

MR. BARANOWSKY: Yes.

MR. DUBE: I'd say that's true.

MR. BARANOWSKY: Right. And I mean don't forget we went to all these plants earlier and did some benchmarking of the SPAR models and we modified them to reflect the as-designed, as-operated plant. And put our standard SPAR modeling assumptions in there.

And in many cases, it matches up --

DR. ROSEN: Yes.

MR. BARANOWSKY: -- or is pretty close.

DR. ROSEN: Okay. Well it's --

MR. BARANOWSKY: In some cases it's not. And these are the ones that aren't.

DR. ROSEN: It's not unsuspected.

MR. BARANOWSKY: Right.

DR. ROSEN: It just makes sense to me now that I know the answer.

MR. BARANOWSKY: So we know there are issues, a few that need to be resolved. We'd either change the SPAR models or they change their plant model. And then we have good agreement.

DR. ROSEN: But that's mostly in the large and the medium?

MR. BARANOWSKY: The large are the ones I would look at. The medium have a very small chance of having an impact.

MR. DUBE: Yes. Now obviously this is just a subset of all the models. It's 11 SPAR models versus the 70 total. So it's a little bit difficult to extrapolate to the rest of the industry.

But if we could do it, I think one finds back to the conclusion I made before which is really major differences in success criteria, major differences in important initiating event frequencies, and the factors of two and three in basic event probabilities which is everyone generally knows is pretty much the norm within the scatter -- don't seem to have an effect.

MR. BARANOWSKY: Actually we did pretty much do what Don's talking about in another program, in the SPAR development program we went and looked at the results of all the other benchmarks that we did. And we made a tabulation. It's very consistent with what we found here and what Don said.

And we're now structuring our enhanced SPAR development work to reflect our understanding of the significant drivers indifferences and numbers of plants that might be involved so that we can get the biggest bang for our buck in resolving these things as soon as possible.

MR. DUBE: The next slide, slide 15, looks at a recent task that we just completed. And it has to do with identifying the system boundary.

And there was a reason for doing this. And that is the concern that, you know, is it absolutely necessary to do a 100-percent inspection of the system boundary that's within scope of the MSPI?

And we wanted to look at the effect of what if a valve was missed because the guidelines for determining the system boundary in the MSPI says all diesels generators and all pumps have to be included. So -- and those generally are the most risk-important ones.

So where we were concerned is what if, for whatever reason, a valve or valves were omitted from the scope of the MSPI system boundary? And what we found is in order for it to have a significant effect on the MSPI, the valve would have had to been inappropriately omitted, the valve would have had to have been a high-risk importance value, and the valve head would have to have a failure rate much higher than the industry norm to impact the MSPI.

And the consequence of omission would be the underestimation of the MSPI, as I said, but the valve would still be subject to the inspection process and at least as currently formulated, an SDP evaluation of the performance efficiencies.

So this was a task, as I said, that we did because we wanted to have an idea of how critical was it to inspect? Was it important to do a 100-percent inspection as was done during the temporary instruction for the pilot program? And our conclusion is no. There's a high degree of robustness.

And the next graph is a cumulative complimentary distribution function of all the delta MSPIs for all 509 valves in the pilot program using a Latin Hypercube simulation. This graph is using a nominal failure rate but we also have it assuming a failure rate five times greater.

And then in this simulation, we took the 95^{th} percentile delta MSPI for each valve. So it's already somewhat conservative. It's not totally bounding but it's an upper level.

And you find -- one finds that if you look at the -- where it crosses the axis here, 98 percent -- 99 percent would have less than a 10 to the minus 8 impact on MSPI. And only a handful of valves would potentially impact the MSPI. And they would have to also have a failure significantly beyond the nominal failure rate.

So it gives us a warm feeling that there is a certain degree of robustness that maybe there is ways of doing the inspection that doesn't require 100 percent verification. In much the same way when a licensee submits a LOCA analysis model, we don't inspect 100 percent of the volumes and the surface areas in the computer code.

And we've expanded this now. We're also looking at the impact of missing a failure or over-estimating demands. And we're seeing that there is a certain degree of robustness as well.

The final topic and it's something we'll spend some time on is an attempt to compare the MSPI, SDP, and safety system unavailability results to the extent possible. And we recognize it's fundamentally different approaches. I mean MSPI measures a statistically valid risk informed change in performance over a three-year rolling interval.

The SSU directly accounts for unavailability but doesn't account for unreliability. And it uses fault exposure time as a surrogate for that. But it also uses a three-year rolling interval.

And the SDP measures short-term risk significance of a failure or condition associated with a performance deficiency. But what it really does, though, is if there is a failure in a particular -- that it is evaluating, it basically if you look at the true mathematical formulation, it uses only a denominator of one year in the calculation of core damage frequency change or change in core damage probability.

We compared the results for all 77 component failures to the extent possible. And all 77 are in that report.

We found that all the non-green safety system unavailabilities were driven by fault exposure hours without exception. And in one case because the T/2 assumption -- because of a T/2 assumption and, in fact, in the current ROP guideline, the SSU -- because of the questions about the validity of T/2, is now excluded from the MSP -- from the SSU calculation.

The T/2 assumption has to do with how one approximates an exponential function by a Taylor series expansion. And there's terms that go T/2 and some higher order terms.

And T/2 works fine for most situations. But when one is looking at a very short time horizon, T/2 can give dramatically erroneous results. And it's a mathematical simplification that, quite frankly, doesn't work.

MR. BARANOWSKY: Well, it can indicate the expected time that equipment is unavailable. When one uses T/2 along with a single year to evaluate unavailability, you can see dramatic swings in unavailability in years where there aren't failures versus the years where there are failures. And, therefore, one's not sure whether they are measuring changes in unavailability or just a normal fluctuation set would occur as you pick intervals short in comparison to the mean time between failures.

I mean that approximation of unreliability breaks down pretty badly when you go to mean time between failures much longer than the observation period. You can see that mathematically.

MR. DUBE: Yes, and we'll show a curve -- a chart on it in a few minutes.

The SDP non-green findings for single failure were often driven by a short assessment period. I said that -- less than a year -- with, quite frankly, insufficient data to measure statistically valid change in performance.

Now I'm not questioning the fact that it is a way of getting to the fundamental heart of whether a deficiency in performance is high, low, medium, or what have you, in risk significance. But as a measure of statistically valid system performance, I think there's some questions.

MR. BARANOWSKY: Okay, I'd just like to supplement that a little bit. I want to make sure that we're not getting the wrong impression here because I had something to do with the SDP process in the early stages being implemented in the Iraqi Oversight Program.

What it does is it tells you at least relatively if not in absolute sense how significant any condition is, okay? It doesn't tell you whether there's been a change in performance. It just tells you something happened and it's significant or not. It could be an expected thing that occurred or an unexpected thing.

And so it has a strength in that it gives a relative importance to whatever the condition that's occurred versus other conditions. And it doesn't have a strength in terms of determining if there have been actual changes in performance.

Now I know this because that process is designed around the action sequence precursor program methodology which we've used for many years and is essentially the same thing. And if you look at any conclusions we've ever drawn with action sequence precursors, you'll notice that we look at multi years of data in order to make any conclusion about whether precursors are arriving or not arriving at a higher or a lower rate.

It's hard to discern changes in performance when you look at one little incident. And I don't think we've ever done that. So it's a good measure of the significance of a performance deficiency. And it has some difficulty in discerning whether or not that performance deficiency is a change in performance or whether it's just the inherent level of performance.

MR. DUBE: All of the -- oh thanks, Pat -- all of the MSPI white or near-white indicators usually involve multiple failures and measurable, significant unavailability that provided a high degree of confidence of adverse change in system performance.

We never saw a situation where it was just one or the other. It always involved significant impact on unreliability, significant impact on unavailability.

We want to see if the MSPI would capture as many if not more reliability/availability performance degradations than SSU and the SDP combined. And a task was undertaken within the Branch. We did a historical review of all 1,659 SDP findings and 5,157 SSU quarterly indications over a 3.25 year period. That was it. We had to cut it off somewhere and it was last summer.

Only 0.5 percent of SSU indicators have been non-green in those 3.25 years. The MSPI results as well as the simulation indicate we would expect, using the MSPI, about 3 percent. It varies between 2.5 percent and 3.5 percent but a nominal number is around 3 percent.

We found that in this time frame, there was an average of four non-green SDP findings per year for the mitigating systems related to actual single failure. That is -- recall that one of the criteria is that if there is a failure, it would have to be detectable during normal surveillance.

If it was a degraded condition or a failure that could not be detected during normal surveillance so that it had a long fault exposure time, it would not be within the scope of MSPI and one would fall back to the SDP.

In the pilot program, there were two white indicators out of 160 systems and three near-white indicators. When I say three near white, they were near white for a number of reasons. First the data collection stopped in the first quarter of 2003. And since this is a three-year rolling indicator, it can't project into the future. And it's possible that one or more of those three near whites would eventually become white.

And I'll show examples. But there were at least two of them where one more failure in one and one-half to two years going into the future would turn it white. So it is true that there's two white indicators. It's also equally true that there are three very close to the threshold near white and only because we stopped data collection will we never know, I guess, whether they'll turn white.

DR. LEITCH: Can I paraphrase here to make sure I'm understanding what you're saying?

Using the SSU process, you got about 25 -- half of a percent of 57 -- or 5,157 --

MR. DUBE: That's for all the plants though.

DR. LEITCH: -- in all the plants in three years. So I'm thinking about something like seven per year, something like that. Is that right?

MR. DUBE: There's like 400-some odd indicators -- it's a handful, yes. It's about -- it's something like that.

DR. LEITCH: Does the half a percent refer to half a percent of 5,157?

MR. DUBE: Yes.

DR. LEITCH: So half a percent is about 25 and you got that over three and one-half, four years or so? So you're talking about seven per year? Something like that, right?

MR. DUBE: Seven to eight --

DR. LEITCH: Right.

MR. DUBE: -- yes, something like that. Then you use --

DR. LEITCH: Yes, go ahead. That's with the SS -- that's with the current SSU --

MR. DUBE: Right.

DR. LEITCH: -- indicators. Now you looked at -- the next bullet down is with the SDP process?

MR. DUBE: Add another four to that eight.

DR. LEITCH: Another four? That's what I'm not clear about. Is that -- is that --

MR. DUBE: Beyond that --

DR. LEITCH: -- beyond that eight?

MR. DUBE: -- beyond that eight, yes.

DR. LEITCH: So there's 12? So the

current --

MR. DUBE: Exactly right.

DR. LEITCH: -- the current on process -- on average the number of non-green SDP and SSU is about --

MR. DUBE: That's right --

DR. LEITCH: -- 12 per year.

MR. DUBE: -- 12 per year for the whole industry. And the MSPI based on the analysis and simulation we get -- there's uncertainty between 8 and 18 -- kind of the upper and lower bounds. They're essentially the same number.

DR. LEITCH: Now the same numbers but are --

MR. DUBE: Are there differences?

DR. LEITCH: -- they the same events?

MR. DUBE: Not always, no.

DR. LEITCH: Not always.

MR. DUBE: And I'll touch upon them --

DR. LEITCH: I guess what I'm trying to visualize -- you've got two sets of events, some with the current system, some with the proposed new system. You know, to what extent do they overlap --

MR. DUBE: Right.

DR. LEITCH: -- since sometimes they do

and --

MR. DUBE: Sometimes they do and sometimes they don't.

DR. LEITCH: -- sometimes they don't.

MR. BARANOWSKY: That's a good point. So there were two things that we were looking at. One was are we going to get a lot more or a lot fewer hits with this indicator. We want to understand the indicator. And so we've got some idea on that.

DR. LEITCH: So as far as I'm concerned --

MR. BARANOWSKY: The second thing is --

DR. LEITCH: -- the numbers are about the same.

MR. BARANOWSKY: -- the second thing is are they going to be different --

DR. LEITCH: Yes.

MR. BARANOWSKY: -- and where are they going to be different and why are they going to be different. Now I expect them to be different -- otherwise we wouldn't have done this.

MR. DUBE: Yes, right.

MR. BARANOWSKY: We think there are some problems with the current indicator and we're trying to fix it. So anyone who thinks that you're going to get the same results, I'd say why did we spend a million bucks and four years of effort? We did it because there were problems identified and we addressed them.

And now, the next chart, if this is the appropriate time, we can go through -- what is it -- about five or four or six specific cases --

DR. LEITCH: See if it's giving you --

MR. BARANOWSKY: -- in which we said why are we getting a difference? Why are they agreeing? What's going on here?

DR. LEITCH: Okay, yes.

MR. DUBE: Okay, so this is a lot of the meat of it, of the discussion. The color is significant here. White means a white indicator. Or finding green means green. There's no yellow and there's no red. And the gray means indeterminate, of course.

Now again in the report, all 77 are discussed and then more cases than this are discussed on an individual, case by case basis, the more important ones. But we narrowed it down here for the purpose of this to the really most significant, most important comparisons. And there are differences.

Braidwood 1, there were three failures of the aux feedwater diesel pump. The MSPI was at 2E to the minus 6. These are rounded numbers. The SDP out of those three failures had one green finding. If you look at the inspection report, there was a green. And the SSU was two and a half percent unavailability over this three-year time frame, which would be white.

And as a comment, the MSPI white comes from a combination of unreliability and unavailability. So here's a situation where the MSPI is white. It kind of matches with the SSU but the greens -- the SDP, the one case, it was a green.

And again the MSPI, one failure wouldn't have turned it white. Two failures didn't turn it white. Three did. But it was also in combination with a significant contribution of unavailability.

One can draw their own conclusions from this case but it did -- I believe the MSPI did what it was intended to do in this particular case.

DR. LEITCH: Could you say something about how you reached that conclusion? I'm just not sure. How does it highlight that? In other words, you get a number like 2E to the minus 6 --

MR. DUBE: Well, it's reflected in the fact that in terms of unreliability, one would -- one can expect failures of aux feed pumps. But what one found is that there were more failures than expectation on the aux feedwater pumps, the diesel pump. And the diesel-driven pump has relatively high-risk importance.

And then in addition, there was substantial unavailability. So it was the combination of contribution unreliability and a contribution unavailability that put it into the white.

DR. LEITCH: The white, yes, yes. Whereas with the SDP, it would have been green. And I guess what you're saying is okay, that's a difference. But that's kind of an explainable difference. And the difference we would have hoped would surface, right?

MR. DUBE: Well, the main difference is that MSPI is a cumulative, rolling, three-year average. So it doesn't just look at one failure. It doesn't just look at the second failure. It doesn't look just solely at the third. It aggregates them.

And so one failure in three years wouldn't turn it white. Two failures in three years wouldn't turn it white. But that third failure doesn't.

DR. LEITCH: Yes.

MR. DUBE: Whereas the SDP in its current form looks at it individually.

DR. LEITCH: Yes.

MR. DUBE: So it looks at the first

failure, it's green. The second failure is green. Third failure is green. And it looks at it in isolation whereas the MSPI aggregates it.

DR. LEITCH: Yes.

MR. DUBE: That is a fundamental difference.

DR. LEITCH: Good.

MR. DUBE: And I'm not -- you know, I just want to point that out.

On Hope Creek, there were three failures --

DR. SHACK: Just coming back to that one.

MR. DUBE: Yes?

DR. SHACK: But apparently the unavailability was high enough to through you over the SSU. So --

MR. DUBE: Yes.

DR. SHACK: -- I'm not sure how you're conclusion that it's a combination --

MR. DUBE: Well, that unavailability -- this all came from fault exposure time --

DR. SHACK: Oh, this is fault exposure time?

MR. BARANOWSKY: And a generic model with generic thresholds --

MR. DUBE: That was --

MR. BARANOWSKY: -- that may or may not be applicable to the specific Hope Creek case. I think --

MR. DUBE: Braidwood case.

MR. BARANOWSKY: You're comparing an apple and an orange and we're just showing you that in this case, they both look alike.

MR. DUBE: As I mentioned on the previous slide, every SSU white, without exception, was white because of its large fault exposure time, hundreds of hours, many hundreds of hours.

The Hope Creek -- there were three failures of high pressure coolant injection MOVs. MSPI was rounded to 10 -- it was above 10 to the minus 6 but Pat says don't show more than one significant figure on these so we just showed it rounded. But it was above 10 to the minus 6. There was no SDP evaluations, reports that we found on any of those three failures. And the SSU was green, 1.7 percent versus a generic threshold of 4 percent.

And again the MSPI white came about from a combination of unreliability and unavailability.

Palo Verde was one failure of a motor driven aux feed pump. The frontstop was applied and made it 4E to the minus 7. Otherwise, it would have been white. There was no SDP done and the SSU was green, 0.5 percent. And as I said, we did sensitivity studies and one more failure over a three-year rolling period would result in a white for Palo Verde.

We could have a whole discussion on the frontstop but this -- it did perform as intended and that one failure would not result in a white indication.

The San Onofre-2, there were six failures of the salt water pumps. And if you recall last time, or if you read through the report, we've put something called a backstop in. A backstop is a way of ensuring -- and it's more performance oriented than risk oriented -- a way of ensuring that if there is a statistically significant departure of the observed failure rate beyond what one would reasonably expect -- and if we want to get into detail, I'll ask Dr. Atwood to get into it -- then regardless of what the MSPI says, we would call it white.

And it's a function of what is the expected number of failures of that component type over the three-year period versus how many were observed. And there's a linear regression that's been drawn.

For this particular component type, and the observed number of demands that there is, the backstop was seven. Which meant if it hit -- if there were seven failures, regardless of the risk importance, we would call it white.

Well, there were only six, so obviously it didn't hit the white threshold but again we stopped the data collection. So it would have been possible for this to turn white but we didn't pursue it.

There was no SDP and because this is a support system, there was no equivalent SSU. It's not applicable.

DR. LEITCH: Okay, I can see from a performance point of view the backstop. Now come back and explain to me the frontstop again -- the one failure. Yes, but why throw out the one failure?

MR. BARANOWSKY: It's not thrown out.

DR. LEITCH: Well I mean --

MR. BARANOWSKY: It's a statistically indeterminate result. In other words, it has as much chance of being an error as it does not being an error. And the philosophy that we applied was that we should have reasonable assurance that there has been a change in performance.

And that the licensee should have some

opportunity to take corrective actions before that change in performance occurs. Oh, we could do the maintenance rule. We don't need a licensee.

I can tell you all the components -- all -- that one failure will kick you up over 10 to the minus 6. I don't need any calculations. I can precalculate them just by knowing their importance and the existing CDF at the plant.

MR. DUBE: Well, it was proposed to address the issue of false positive. And what we found that if baseline performance is near the 10 to the minus 6 threshold, there is a high probability, a high likelihood that the indicator indicates white but performance may be green.

And it's because there is a distribution. We're using mean values here on the MSPI but there's really an uncertainty in the distribution. And there could be tens of percent probability that it's really green.

And it's kind of analogous to the situation where let's say in a particular neighborhood or cul de sac, the incidence of a rare disease occurs. And if one takes one incidence and divide it by a small population base, one might conclude that the incidence of that illness is 10 or 100 or maybe even a 1,000 times the normal expectation.

But any epidemiologist or biostatistician will tell you that's statistically insignificant. One can't draw that conclusion. The frontstop was put on for that particular purpose.

DR. SHACK: Okay so any failure could throw you into an SDP if it was significant enough. But you're arguing that it doesn't really tell you about performance because --

MR. DUBE: Right.

MR. BARANOWSKY: I don't know -- well no SDP was done on Palo Verde

DR. SHACK: Right.

MR. BARANOWSKY: Or yes it was -- no, not done. It may or may not have, I don't know.

DR. ROSEN: But in layman's terms, it's a way of dealing with the fluke, the statistic --

MR. DUBE: Right.

DR. ROSEN: -- the thing that happens and nobody expects it and it doesn't really tell you anything about the performance.

MR. BARANOWSKY: Well, you're not sure. You know -- it could be due to some performance deficiency --

DR. ROSEN: Well, it could be --

MR. BARANOWSKY: -- some poor maintenance.

DR. ROSEN: It could be.

MR. BARANOWSKY: You just don't know whether the performance of the plant is trending downward.

DR. ROSEN: It could be but --

MR. BARANOWSKY: Actually you don't know.

DR. ROSEN: -- you can't use it this way because it is just as well likely it could not be. It could just --

MR. BARANOWSKY: Yes, but that doesn't mean it's not a risk-significant failure.

DR. ROSEN: Right.

MR. BARANOWSKY: So there is a little bit of brain-twisting you have to do in thinking about this.

DR. SHACK: But can't you go back to your statistics to tell you whether a failure was extremely unlikely? I mean the epidemiologist would know that, you know, the one failure is something that meant absolutely nothing. In this case, couldn't I know whether it was or was not likely to have a failure from prior -- I mean the frontstop seems like an absolutely rigid rule.

MR. BARANOWSKY: Well, actually it's based on doing statistical analysis.

DR. SHACK: Okay.

MR. BARANOWSKY: Okay? And I think Don has a following chart which shows sort of visually the -- is that the next one?

MR. DUBE: Maybe not.

MR. BARANOWSKY: Oh, yes. Like for instance, we went back on at least one or two of these and we said what has been the performance over some period of time on this thing? Is this -- because we only captured what -- a year or so of data here? So let's go back and see whether this is a fluke or not.

And if you see something that's failed about once every five or six years, well, perfect performance for four or five years, one failure, bad performance. Perfect performance for four, five, six years, then one failure, bad performance.

It doesn't mean it's not risk significant. It just means that that's the performance.

MR. DUBE: Dr. Atwood, is there anything you want to add along these lines?

DR. ATWOOD: I don't think so.

MR. DUBE: Okay.

DR. ATWOOD: Unless you want me to address a particular point.

MR. DUBE: Okay. The next case here -- and one -- this is a challenging one -- this is a tough one. And there is an inconsistency here. And a big difference.

There were four failures of the emergency diesel generators in the third quarter of 2002. The MSPI is 8E to minus 7. One additional failure through the second quarter of 2005 -- this is one of those near whites where you start collecting data so that it's really indeterminate -- or a net addition of 40 hours of diesel general availability would result in white. So it is a borderline case whereas the SDP was at 9E to minus 6, clearly a white, a high white. And the SSU was green, 1.5 percent.

What really kind of concerns us is the fact that there were four failures in one quarter and the MSPI, as it is currently formulated, does -- would not catch that. And because of that -- and this was -- we kind of discovered this in the last really couple months, you know, we're assessing whether one would add something called a short-term backstop which is like the backstop but the backstop is over a three-year period.

This would look at significant number of failures in a short period of time that like the regular backstop, identifies a statistically significant deviation from the norm where the one should not call it white regardless of risk importance and risk indication.

We haven't really assessed all of it --

MR. BARANOWSKY: But to go back to it -- and you said, Dr. Shack, this is -- we've got some methodology that we developed for the initiating event indicators for -- that we discussed here on another program and which you used some prediction interval techniques to look at short-term deviations which we could apply here.

And correct me if I'm wrong, Dr. Atwood, who has actually developed it -- and if we did that, we could detect short-term deviations fairly quickly. And that's why we've been looking at those things in another program. But they have applicability here.

Would you like to add to that?

DR. ATWOOD: I want to go back to the Palo Verde.

MR. BARANOWSKY: Okay.

PARTICIPANT: Use the microphone.

DR. ATWOOD: Corwin Atwood, Statwood Consulting, contractor for NRC.

The question was asked what if that one

failure at Palo Verde were extremely unlikely, would the frontstop still be applicable? I think then the backstop would because the backstop says what if you get the number of failures that are extremely unlikely.

Now when we calculated our backstops, we always needed four or more failures. But I think that's how we would pick it up.

MR. BARANOWSKY: Okay.

MR. DUBE: And the last case is Millstone-2, there was one failure of a turbine-driven aux feed pump. The MSPI actually is a minus 4E to minus 7, a green. And it's for a reason because it's a system indicator and there are three trains of aux feedwater, two motor driven and the steam driven.

The motor-driven pumps have much better than baseline performance, much better lower unavailability than anticipated, lower unreliability than anticipated. And they more than compensated for the turbine-driven pump that had one failure and it had an unavailability at baseline.

So you have two trains, much better than the norm, one right at the norm, and the MSPI basically is a system indicator. And so we understand that and we think it makes sense. But the SDP was a white, it was a 3E to minus 6. And SSU was a white again because of all -- because of fault exposure time at 2.7 percent.

So these are really the major comparisons and the major differences. And yes, it is fair to say that they're measuring different things. But I think it's fair to say that in almost -- in all cases, the white or the near whites from the MSPI are combinations of, you know, contribution to unavailability and contribution to unreliability in combination.

MR. BARANOWSKY: Don, there was one other thing that -- or maybe two other things. When you did the simulations, did you come up with any expectations for yellow indications? And also maybe you might want to explain how the front stops, et cetera, work with regard to high-risk significance indications that might be yellow in effect.

MR. DUBE: I don't think the simulations found any yellow. But the frontstop would only be applied at the 10 to a minus 6 threshold. If a failure were to put one into yellow or higher -- yellow or red, the frontstop would not be applied. So it's only at the least risk-significant level of 10 to minus 6.

DR. ROSEN: Now this last row is a case of what I think Bill Shack was talking about earlier. It tends to -- because it is a system indicator, you end up not getting for the resolution what you would be from just the --

MR. DUBE: SDP?

DR. ROSEN: -- the SDP.

MR. DUBE: You're right.

DR. ROSEN: Is that right?

MR. DUBE: Yes.

DR. ROSEN: It tends to merge them? And in this case, it over merges them and puts it negative actually.

MR. DUBE: Right.

MR. BARANOWSKY: Well, I don't know about over merging. What it really says -- I mean you have to understand what the minus means. The minus means performance has approved overall. If you want to know what it means in terms of risk, it means I have less change of having a core damage accident because of compensating reliability and unavailability considerations with respect to the rest of the system.

MR. DUBE: In the last three years.

MR. BARANOWSKY: So one train is not performing as well and two other trains are performing much better.

Now maintenance rule should catch that one train and force the licensee to go do something with it.

DR. ROSEN: But in this case, that's not true because the train that is performing worse is performing at the baseline, right?

MR. DUBE: The unavailability was.

DR. ROSEN: The unavailability. But the unreliability was worse?

MR. DUBE: Actually it was better because this only takes a three year rolling time frame but the last failure was over a decade ago of the turbine-driven pump. So if one had a ten-year rolling indicator, it would have been better than baseline. But we used a three-year baseline, a three-year measurement period.

DR. ROSEN: I'm losing my point. But I -- what I'm really seeing here is we've got a system, Millstone-2 aux feedwater system with I think three pumps --

MR. DUBE: Right.

DR. ROSEN: -- two of which are motor driven, one of is steam driven.

MR. DUBE: Right.

DR. ROSEN: And the steam-driven pump, the turbine-driven pump, is at the baseline for unavailability, which means it's performance is nominal, is that right?

MR. DUBE: For unavailability, correct, it was.

DR. ROSEN: And the motor-driven pumps are better than nominal?

MR. DUBE: Correct.

DR. ROSEN: So in aggregate, this is better --

MR. DUBE: Yes.

DR. ROSEN: -- than your nominal value?

MR. DUBE: Correct.

MR. BARANOWSKY: And it also --

DR. ROSEN: And that's what the minus means?

MR. DUBE: Right.

MR. BARANOWSKY: And it just so happens that the stage in blackout action sequences at this plant are not dominant with respect to importance in the auxiliary feedwater system. If they were, then you would see a different result --

MR. DUBE: You might have a different conclusion.

MR. BARANOWSKY: -- that's what makes this plant specific. You put that failure in another plant --

MR. DUBE: It could be --

MR. BARANOWSKY: -- where the station blackout sequences are dominant, and then the Fussell-Vesely factors change and you end up with a different --

MR. DUBE: It could be white.

DR. ROSEN: Well what's interesting to me about all this is that this formulation, the MSPI is extraordinarily rich in terms of information. If you question it, it gives you something to look at.

Then you say, well, what does that mean? And once you start asking that kind of question, you get answers that have some meaning. And to me that's better than simply an SDP that says -- or an SSU that says two percent. Is that good? I don't know. Compared to two and one-half, it's okay.

Well, was two and one-half any good? I don't know. You know you don't get anywhere when you talk about SSU really. You just think you might be getting someplace but you're not.

I like MSPI because it's -- as long as you use it intelligently, as long as you say okay, what does that minus 4E to the minus 7 mean? You get a lot of answers that have some meaning.

MR. DUBE: Well, it brings up the features of it's a trained system indicator -- I mean a system indicator and it's measuring against a baseline performance and over a three-year period. And I think it's done with some, you know, anomalies that we might have to tweak. But for the most part, it gives you results that you understand and you can explain.

DR. SHACK: Well, your frontstops and backstops address a lot of my concern because you're going to sort of catch performance problems even if they're not risk significant because they're going to bump into those backstops. And your short-term backstop would even help that problem a little bit more.

MR. DUBE: I think so. In fact, you know, Dr. Atwood did an analysis. And if we were to have a short-term backstop, it turns out -- believe me, we didn't even look at Salem-1.

It came out to four failures in any two quarters, two sequential quarters, so a total of four. And that's probably what we would use. But we haven't really discussed this at any public forum. But that's -- we would use that for any component type. It turned out to be a pretty, you know, constant number.

DR. SHACK: I'm sort of amazed you can get these universal results if the backstop is 7, you know.

MR. DUBE: The backstop actually is a linear regression.

DR. SHACK: It's like pi, right? Or E equals MC^2 -- how simple could it be?

DR. SIEBER: Would this be an appropriate place for us to take a break?

MR. DUBE: Sure.

MR. BARANOWSKY: I think we're ready.

DR. SIEBER: Okay. Why don't we come back at 25 after ten.

(Whereupon, the foregoing matter went off the record at 10:00 a.m. and went back on the record at 10:22 a.m.)

MR. DUBE: I guess we're all back or on the way back. Were there any questions on this slide 19? If not, I'll just continue on. We're almost done with the technical presentation although I do have some backup slides on two questions that were asked.

The next slide shows in a kind of generic sense, although it's actual data from one case, how the application of fault exposure time to single component failure can lead to a volatile indication of performance otherwise at industry norm, that is false positive. And if one just took a turbine-driven aux feed pump that was tested monthly and assumed a 14-day fault exposure time, and if it had nominal importance measures and Birnbaums, and the nominal probability failure to start of 10 to the minus 2, which is what we found to be norm for steam-driven aux feed pumps.

One would calculate a mean time to failure of six years. That includes demands from additional operations. So one could see using the fault exposure time concept how one would go along and have five years of good performance in green and then guarantee that every sixth year or roughly every six years on average, to be in the white range -- I mean inevitably.

And if one were to use and apply the fault exposure time in that sense, it can lead to a false positive indication because the pump is performing exactly at the industry norm, no better, no worse, and yet every sixth year, one would get a white indication or white finding doing nothing more than the current -- using the current process of looking at an ASP-type of calculation.

And in that sense, you know, that is a big significant difference between the MSPI which aggregates it over three years -- and the three years was chosen for a specific reason. As part of the NUREG-1753, a spectrum of ranges were looked at.

We looked at as short as one year and even longer. One year resulted in too volatile, five year resulted in too much of a delay and lag, and three years seemed to be optimum or just right.

And I just wanted to bring this up because it is a fundamentally -- you know, a fundamental difference between a SDP/ASP-type of evaluation and the MSPI which aggregates it over three years.

And in that sense, it leads me into the next slide whereas one of the qualities of MSPI is it addresses both false positive and false negative concerns. The frontstop, in conjunction with the backstop and CNI, constrained noninformative prior, effectively constrained the minimum and maximum failures to white.

The backstop prevents false negatives in the sense that one -- you know, if you use just the algorithm, one would calculate 10s if not 50 or 100 failures to white. The backstop says we're going to constrain that. And if there's a statistically significant departure from expectation, we're going to call it white regardless of risk.

The frontstop constrains it at the other end to prevent false positive and the three combined, constrained noninformative prior, which was a fallout from the NUREG-1753, it was found to be the best that we looked at at the time. The three combined seemed to give reasonable results.

I think that all those words summarize all those bullets there. And the other important thing to point out is that latent faults, that is a condition that can't be discovered by normal surveillance that can result in a large fault exposure time generally are those that result in large -- potentially large risks or multiple concurrent failures which -- where the synergistic effect of multiple failures at the same time result in high risk, those would continue to be evaluated by SDP. It's important to point out.

So the MSPI if it were to supplement or supplant and substitute for SDP would only do it for those situations where it was no concurrent failure and a single failure than can be detectable by normal surveillance. Otherwise, the SDP would be the approach to use -- the method to use.

And then finally, MSPI --

MR. BARANOWSKY: Let me just --

MR. DUBE: Yes?

MR. BARANOWSKY: -- I wanted to just back up -- back again to that bottom bullet and make sure that we were clear about the importance of these latent faults and multiple component failures. Those are the high risk events that the history of operating experience shows us are important, okay?

So we in no way want to miss those. And we want to have a fairly quick and sharp identification of the significance of those events and -- well, that's one of the reasons why we have not put that kind of thing into the MSPI.

They are the same kinds of events that have been found over the years to be important through the action sequence precursor program. They're -- they can be at the higher risk level with respect to core damage frequency whereas single failures generally tend to be at the lower risk end of the core damage probability distribution.

So the highest risk concerns should be able to be identified promptly with an appropriate methodology.

DR. ROSEN: Pat, let me -- while you're on this slide -- make a point --

MR. BARANOWSKY: Okay.

DR. ROSEN: -- about risk communication. And that is in a program like this which I believe has a lot of important and useful features and if it goes forward in the Agency and becomes part of the ROP, it's got to be explained to the public and to lay people at large.

And using terms like constrained noninformative prior and frontstops and backstops is not going to be helpful. So you've got to figure out a way to put this in lay terms that helps, you know, people who are not PRA nerds like yours truly and others understand what we're talking about.

MR. BARANOWSKY: Yes, I appreciate what you're saying and some aspect of this, if we move forward, needs to be put into the simpler terms and it can be done. But the technical details that we're talking about now need to be discussed at the appropriate level.

And I can't, in a meeting like this when we're getting into technical issues, fix that situation but I appreciate it.

DR. ROSEN: Well, I'm glad you didn't here because --

MR. BARANOWSKY: Yes.

DR. ROSEN: -- we -- some of us have got to tolerate this sort of thing and it's okay.

MR. BARANOWSKY: Right.

DR. ROSEN: But I think when you go more broad with this, if we go more broad with it, you need to be thinking about risk communication and with a different set of terms --

MR. BARANOWSKY: We'd have to --

DR. ROSEN: -- to describe the same things.

MR. BARANOWSKY: -- write a sort of an executive summary for the laymen if you will. That's a good point. And we're learning about risk communication so I'm looking for the experts who can help me write that. And it turns out that I'm looking in a mirror.

DR. SIEBER: I guess I have -- I would agree with what Steve is saying. And to me that's the most difficult part of using this kind of an indicator is that it's very complex and it has a lot of features to sort of adjust it so that it works right, which the average member of the public may not be able to fathom properly.

MR. BARANOWSKY: Yes.

DR. SIEBER: And I think explaining it in terms that the general public would understand is going to be an extremely difficult job. It will probably be as difficult as developing the indicator was in the first place.

DR. KRESS: But on a different note, if it's the right thing to do, then don't let that stop you just because it's hard to explain.

MR. BARANOWSKY: Understood. There are some complexities associated with the development of the indicator but remember the bulk of what we are doing is looking at how this actually works and what its outcomes and unintended and intended consequences are so that we can, in the end, summarize very simply what the indicator is and how it works.

And leave the technical report on a shelf for the eggheads like us to look at. And have a much shorter version of what this is later when it's appropriate. And that might be sooner than later but --

DR. KRESS: You know I have a feeling that the general public will understand. You're looking at important systems and components.

MR. BARANOWSKY: Yes.

DR. KRESS: And if they're out of service or fail too often, then it's a measure of performance and you go from a green to a white if these things are too significant. I mean I don't think you have to get into the details of --

MR. BARANOWSKY: Yes.

DR. KRESS: -- of the other stuff. They'll understand that.

MR. BARANOWSKY: We haven't tried it yet on this project but you know the NRC has recently published some risk communications guidelines and we are using it on some action sequence precursor results that have come out.

And I think we're getting the handle on it. And ultimately this is -- it could be applied here, too, and we just haven't done it yet.

DR. SIEBER: Will you have done that to any extent by the time you issue your final report in September?

MR. BARANOWSKY: Maybe we'll take a cut at that.

DR. SIEBER: I --

MR. BARANOWSKY: Yes.

DR. SIEBER: -- I think that it's enough of an integral part of this project that we ought to have some indication when we meet again in September as to how you're going to deal with it.

MR. BARANOWSKY: Okay. Now what we're talking about is not rewriting this so that a school teacher could read it. But taking the executive summary, condensing it into a smaller, succinct version that doesn't have jargon in it that drives lay people crazy. Right?

DR. KRESS: What I'd be interested in is is this going to actually become part of the ROP? Or is this just an academic thing that you looked at this? Or are there actually plans to revise the ROP and include this?

MR. BARANOWSKY: I don't know. And, therefore, we have an exciting speaker that's going to follow me by the name of Bruce Boger who can help you with that.

DR. KRESS: Okay.

MR. DUBE: I think you have the next one.

MR. BARANOWSKY: Okay, the -- let me just wrap this one up, No. 22. We talked about this a little bit earlier. But the MSPI we think is consistent with the maintenance rule, tech specs, and the principles as follows in SECY 99-007, which were the recommendations for the ROP improvements.

Specifically for the maintenance rule, we looked at definitions of failures, demands and unavailable hours.

And actually we had some accommodation between maintenance rule folks and us on how to capture unavailability, especially during shutdown conditions versus at-power, which I think went a long way to helping the bookkeeping, if you will, on collecting unavailability information.

Also the bases are consistent with technical specifications in the maintenance rule, both of which are tolerant of single failures. The technical specifications -- in fact, the regulations themselves are tolerant of single failures. It doesn't say single failures can occur over and over again. But it does say a single failure is tolerated in the technical specifications, including in the risk informed technical specification development.

DR. ROSEN: It's because that's the way the plants are designed.

MR. BARANOWSKY: Well, of course, if they're not, the implications would be to constrain surveillance intervals to such a small time frame that one couldn't possibly have risk above 10 to the minus 6 when looked at over a short period of time.

DR. ROSEN: Well my point is that it doesn't -- the technical specifications of the maintenance rule don't exist in isolation. They exist and are workable because that's the way -- they reflect the way -- the design basis of the plants and the regulations.

MR. BARANOWSKY: Right.

DR. ROSEN: So they're consistent across the board. It's what we call -- we're starting to call coherence here, you know.

MR. BARANOWSKY: Right. Exactly my point. The maintenance rule is meant to be the first line defense in maintaining and detecting performance, especially performance changes. The performance indicators come in next. That's the philosophy in 99-007. And the technical specifications provide an ultimate umbrella or limit on things.

If you look in 99-007, you'll see statements like sufficient margin should be provided in performance thresholds so that there's an opportunity for licensees to take appropriate action. And where practicable, we should use performance indicators and inspections to cover other areas.

And so we've tried to design an indicator around that idea. Of course there is some debate about what is practical and not practical with respect to the MSPI and I think that's still being looked into. But Bruce Boger can tell you more.

So the conclusions. Although we haven't completely finished our work, I would have to say that the MSPI has been really thoroughly tested and evaluated. And, in fact, discussed at many public meetings. Not just here but with stakeholders almost on a monthly basis over the last couple of years.

The indicator clearly addresses the problems that were identified with the current set of PIs. As you've seen from all the sensitivity studies, simulations, and unbelievable number of calculations that were done here, we have a very good understanding of the capabilities, strengths, and limitations of the MSPI.

Some points were raised here that maybe we need to make sure are clear on our understanding of strengths and limitations. We're looking at a few what I call peripheral adjustments -- this adjustment to the short-term capability of detecting deviations from nominal performance.

And also any other comments, by the way, that we might receive either from external or internal stakeholders. Because we're in an internal and external public comment period on the methodology that's captured in this report. And we won't do a final report until we address those comments.

Nonetheless, even in the current configuration, I think we've shown that the MSPI is quite a capable indicator. It's desirable qualities involve its ability to be very plant specific. It has a proper treatment of reliability and availability. It allows balancing and weighs them appropriately.

It captures performance degradation and considers false positive and false negative concerns. And the results are pretty robust as you could see from some of the analyses that we did, sensitivity studies, and the simulations.

I mentioned its consistency with the maintenance rule and tech specs. The PRA adequacy issues that Don went over in some detail are identifiable and potentially manageable. They're not resolved at this point but they're potentially manageable.

Some contribute to significant discrepancies in outcomes but because of the robustness and the limitations and the structure of the MSPI, those smaller ones have virtually no impact on the outcomes. You get the same outcome.

DR. ROSEN: Before you get off that slide, could you say more about this bullet computation is structured and programmable. What did you mean?

MR. BARANOWSKY: Yes, in fact maybe this is where I should use this one back up view graph.

MR. DUBE: I think it also answers your question on the data from earlier.

MR. BARANOWSKY: Yes.

DR. ROSEN: My question was about the industry and the staff's difficulty in handling the data.

MR. BARANOWSKY: Exactly. So this will address that. No, you just have to twist that.

There are a lot of subtleties in the development of the MSPI methodology but when it comes to doing the calculation, one calculates UAI and URI. Those are the two indices, reliability and unavailability index. There are one-time inputs of the core damage frequency, Fussell-Veselys, the baseline unavailability for the plant that come directly from the PRA, okay?

The data that one collects for unavailability, of course, is the number of hours for -- unavailable hours for each MSPI system train when the plant is critical and the number of critical hours, which is clearly straightforward. We're talking about, you know, a dozen trains or something like that.

DR. ROSEN: But this is not new.

MR. BARANOWSKY: This is not new. This is maintenance rule. That's why I'm saying it's consistent with the maintenance rule. They're collecting that information. What we need to do is make sure that the information is collected in a common format. Thus we had the activity at INPO to create the consolidated data entry system so that one collects this information one shot. And I'm going to let NEI say more about the practicality of that.

DR. ROSEN: But I -- the first two lines on this slide, you haven't said one word -- and I'm not being critical, Pat, I'm just pointing it out.

MR. BARANOWSKY: Okay.

DR. ROSEN: You haven't said one word about any new data yet.

MR. BARANOWSKY: There's no new data.

DR. ROSEN: Okay.

MR. BARANOWSKY: The unreliability uses, of course, the core damage frequency, the Fussell-Vesely importance measures related to unreliability, baseline unreliability data, which, by the way, you already have. That's not something new. And the --

MR. DUBE: Mission time.

MR. BARANOWSKY: -- mission time for the component and some specific parameters which are already tabulated and they're not subject to change at this point.

DR. ROSEN: Okay, let's play our game again. Still nothing new, right?

MR. BARANOWSKY: Nothing new there. The data collected quarterly would be the numbers of demands, failures associated with those demands, run hours and failures associated with run hours for approximately I'm going to say 30 to 50 components per plant.

DR. ROSEN: In total?

MR. BARANOWSKY: Yes. Total, for all systems.

DR. ROSEN: Is this all new stuff or is this --

MR. BARANOWSKY: No. This is the same stuff --

DR. ROSEN: Oh.

MR. BARANOWSKY: -- that one would collect for the maintenance rule or to do a PRA.

DR. ROSEN: So still nothing -- still nothing new.

MR. BARANOWSKY: And what one does is

takes the information that I have in the second line of UAI data and URI data, feeds it into the consolidated data entry program, which is I think still in development but it's -- if someone was INPO here they could tell me -- and that would compute the outputs.

So it's -- although there is a sophisticated amount of thinking that went behind the methodology, the elements that one deals with routinely are pretty much the basic things that if you can't do this, you can't do PRA. I'm standing by that. But this is easier than doing PRA.

DR. ROSEN: I'd also say if you can't do this, you can't do maintenance rule.

MR. BARANOWSKY: I don't see what -- this is it. So I don't know if that answers your question but --

DR. ROSEN: No, it does.

DR. SIEBER: Well if you can't do PRA, you'll never get to this, right?

MR. BARANOWSKY: Right. It's sort of a circle. It's like the chicken and the egg.

DR. SIEBER: Right.

MR. BARANOWSKY: Okay. So I went -- let's go to the last -- okay, this is our current plan on what we need to do. The internal comments are due in May. We issued the report to NRR and the regions in the end of February so we've got March, April, May there.

Public comments are due in June. We'll take the comments -- we're presuming that they're the usual modest amount of comments because we've taken this methodology and put it out so often we know the issues and, you know, either we're going to make a few adjustments or we're going to describe why we think what we already have in place is adequate.

We'd be quite interested in any comments from ACRS members. Then we'll try to have a final draft of the MSPI in August, which we would make available to the ACRS.

And hopefully, if you agree, come to the ACRS full committee in September without having another subcommittee at which point we would ask a letter on the MSPI technology or methodology, if you will. And then issue a final report in October of 2004, this year, which is pretty much the schedule we laid out two years ago.

That completes our presentation.

DR. SIEBER: All right. Does any member have any additional questions they'd like to ask at this time?

(No response.)

DR. SIEBER: If not, I guess we can move ahead with the agenda. And ask Bruce Boger from NRR to come forward. Or you can do it from there.

MR. BOGER: Well, it sort of puts me at a disadvantage -- I prefer --

PARTICIPANT: Bruce, take my chair.

DR. SIEBER: We have plenty of chairs.

PARTICIPANT: Do you have any slides, Bruce?

MR. BOGER: I have no slides.

PARTICIPANT: Okay.

DR. SIEBER: Okay.

MR. BOGER: Good morning, my name is Bruce Boger. I'm the Director of the Division of Inspection Program Management at NRR. Among my responsibilities in that position is I'm responsible for the reactor oversight process and I'm also responsible for the determination of regional inspection and assessment resources. So some of that comes together in my division.

And I wanted to have a chance to speak with your this morning to tell you where NRR is headed with respect to finding a replacement for the safety system unavailability or SSUPI. Now we've had a lot of discussions this morning. You know, I think that there are -- a lot of folks are in agreement that the MSPI has several advantages over the SSUPIs. It solves many of the problems.

However, we also think there are some disadvantages to use of MSPI relative to the reactor oversight process. Last month, in March, we discussed those with the Commission.

The Commission has provided us a staff requirements memorandum relative to this. And basically they have encouraged us to continue the pursuit of a risk-informed performance indicator that resolves the issues associated with SSU performance indicator.

They have encouraged us to do so in a timely manner. They advised us that resource considerations should not be a primary consideration in moving forward in a risk-informed way.

They encouraged us to address the lessons learned that we had from the MSPI Pilot Program. And they asked us to continue to involve stakeholders in that process.

Right now, today, there is the Agency Action Review Meeting that's taking place with senior

NRC managers. One of the topics of discussion is the MSPI and the approach. We're planning to talk to the Commission in May about the Agency Action Review Meeting results so, of course, we'll talk about the MSPI at that time also.

DR. SIEBER: Do you have a date for that meeting?

MR. BOGER: The Commission meeting is May 4th.

DR. SIEBER: Oh, okay.

MR. BOGER: We envision a series of meetings as we roll this out with whatever direction we receive. But we see a series of meetings taking place involving internal stakeholders and ultimately external stakeholders in determining the approach on how to move forward.

That, of course, would include the ACRS in that stakeholder evaluations. So that's where we're headed in NRR.

DR. KRESS: You mentioned that you see that the MSPI has a lot of advantages. But also has disadvantages. Do you care to mention what you --

MR. BOGER: I think they --

DR. KRESS: -- perceive those disadvantages to be?

MR. BOGER: -- we could probably have a nice session on that. Those issues are outlined in the Commission paper that we wrote, which is SECY 04-0053. There are several -- there's a discussion on performance indicators and, in specific, the MSPI.

DR. SIEBER: Get me a copy of that.

PARTICIPANT: Yes, we need a copy of that.

MR. BOGER: But we would -- if the --

DR. ROSEN: Could you give us some highlights of what 04-0053 cites?

MR. BOGER: I could read them for you. I'd prefer not to do that. I can tell you that on a very high level, I think the -- we're the implementors. We're the implementors of this technique. And many of the implementors are not satisfied that the way the pilot program or the MSPI Pilot Program was piloted, what existed in that pilot gets us to where we want to be.

DR. SIEBER: Steve, I've asked to have copies of that SECY paper given to us or sent to us.

DR. ROSEN: Is that a fixable problem or do we have to go re-pilot it in your view or what are the implications of that -- the implementors don't like the way the pilot was done.

MR. BOGER: No, the didn't like some

aspects of the MSPI as piloted. For instance, they use of the significance determination process, the use of the frontstop, the availability of PRAs to the public. I mean there were aspects like that that we tried to balance in our decision.

And so our going forward would be learning from that, taking a look at what's -- what the difficulties that we have from the SSU and coming up with something risk informed that replaces the SSU.

DR. ROSEN: So, I'm sorry, I'm trying to write and listen at the same time, Bruce.

MR. BOGER: Yes, sir.

DR. ROSEN: The use of the frontstop, the availability of PRAs to the public and there was one other thing you said that --

MR. BOGER: What did I say -- frontstops --

PARTICIPANT: SDP.

MR. BOGER: Oh, the use of the SDP or not use of the SDP under certain situations.

DR. SHACK: Or whether you would have to do an SDP if you had the single fit -- whether the SDP would replace the MSPI? Or the MSPI would replace an SDP? Is that the issue?

MR. BOGER: Right for single failures.

DR. SIEBER: Now the PRA that you would use in this process is really the SPAR models?

MR. BOGER: Well --

MR. DUBE: No, there would be the licensee's PRA -- plant PRA.

DR. SIEBER: The licensee's PRA? Okay.

MR. BOGER: And there are, you know, underlying concerns about PRA quality and consistency as well.

DR. SIEBER: Right. But you're going to have that problem at this point in time until improvements take place with every risk-informed application that you have.

MR. BOGER: Right. And that would lead us to want to perform inspections or at least do an evaluation of what a licensee was using to perform the MSPI which is resource intensive.

DR. SIEBER: Yes, it is. Okay.

Any other questions?

DR. ROSEN: Let me think about strategy here now. When we get a chance to read 04-0053 sometime soon --

DR. SIEBER: Yes.

DR. ROSEN: -- then will we have a chance to discuss that with any -- with the inspection branch?

DR. SIEBER: We don't any --

DR. ROSEN: We've had the chance to discuss the research reports with the research people but --

DR. SIEBER: Yes, we will not have an opportunity until the September full committee meeting at which time, you know, our thoughts need to be pretty firm.

So we would not at that point in time be in an investigative mood. Because if you're going to produce a letter, you should have done your investigation, got the questions answered, before we start drafting the letter.

So if we want to review what's in 0053, I think that we would have to have some kind of a subcommittee meeting in advance of that. And I suggest that we wait until we get the document and to read it to see, you know, whether we think that that kind of a meeting is necessary.

MS. WESTON: Right because if it contains primarily policy information, we will not be having a subcommittee on that.

DR. SIEBER: Yes. If these are -- if it's resource or management issues --

MS. WESTON: We will not be discussing it.

DR. SIEBER: -- we're -- our comments are on the technical end.

MS. WESTON: Right.

DR. SIEBER: As opposed to management issues.

MS. WESTON: Right.

DR. SIEBER: Okay, any additional questions?

(No response.)

DR. SIEBER: Seeing none, thank you very much, Bruce.

MR. BOGER: You're welcome, thank you.

DR. SIEBER: And lastly we have some public comments from NEI. Tony Pietrangelo will provide those to us. So welcome Tony.

MR. PIETRANGELO: Thank you. I appreciate the opportunity to be here.

My comments will be pretty brief because I think in the earlier presentations, most of the stuff I wanted to cover were already covered. So I'll chose to kind of underscore what are some of the more important aspects of this to us.

Starting with -- well first of all, I just want to compliment the NRC on the technical work that was done in the methodology for MSPI. There was more work done to develop and define the technical basis for MSPI than what was done for the entire ROP over four years ago.

This thing has been pushed, prodded, tested more than any of those other PIs. Okay? So we understand what the potential impacts are. We understand what the limitations are. I think there is a lot to be said for that. So to Pat and Don and the contractors, as well as the regions and NRR folks who participated in this, I think it's an excellent technical report.

We don't have any technical issues. As a matter of fact, there was so much interaction over the last two years with the pilot plants, with the PRA folks providing input to Don and his contractors on this, technical issues were identified, solutions were proposed, pilots provided additional information to look at what the impacts would be.

It was a very collaborative process that went into that technical report that identified what some potential resolutions were. And then we got agreement on what those technical issues were and their resolution.

Again, from our perspective, there are no technical issues left with MSPI. Whatever is in that SECY I think is more -- was alluded to more policy/management/resource-type issues. If there are, in fact, technical issues left, we want to hear about them so we can resolve them.

Okay, let me go back now. Why was the industry interested in this effort from a technical standpoint? And most of this was alluded to in Pat and Don's presentation. We had multiple unavailability definitions in the industry; one for the ROP SSUs, one for the maintenance rule, one for the WANO/INPO indicators, and even some different PRA definitions.

So if you asked the system engineer at the plant what's the unavailability on the RHR system, you know, he said well which answer to you want? Do you want the one for the ROP? Do you want the one for the maintenance rule?

The bookkeeping was mentioned. That is a significant concern for the industry. In fact, the main driver for, you know, the industry effort on this were our CNOs. They were hearing complaints at the plant about data collection and how confusing it was, how resource intensive it was.

MSPI, from our standpoint initially

anyway, was the way to get a single definition for unavailability that could be used to support different things but that was consistent.

The cascading of the support systems was mentioned. The unavailability monitoring during shutdown was mentioned.

One thing that wasn't mentioned was that a lot of the -- in terms of the SSU unavailability definition was the design basis context for unavailability.

We're changing that to a risk-informed context consistent with the maintenance rule as well as PRA. And WANO is willing to accept that. So that's a tremendous change in terms of the context for the unavailability data is collected. And leads to a more consistent risk-informed approach.

There was also inconsistency with -- you had performance criteria for all these systems in the maintenance rule, both unavailability and unreliability, okay, that are risk informed and then you had these generic thresholds based on performance in the SSU.

And because the systems have different risk importances based on the plant-specific designs across the industry, you had kind of a mixed fruit salad of where the performance criteria was versus where the generic performance threshold was for all these different systems across the industry.

We believe that once MSPI is in place, and we hope it will be in place soon, you get rid of that fruit salad and you have a nice alignment of the maintenance rule performance criteria with the threshold -- plant-specific thresholds and MSPI for those systems. So we see a tremendous benefit in terms of the coherence of where those performance criteria and thresholds are.

And I think another big benefit from a technical standpoint with MSPI is that you know up front what are the risk importances of those components in those systems. You know up front. That allows you have a much greater focus on safety within the context of that program.

The data collection was mentioned before. There is no new data. There is some -- a one-time effort to collect these different importance measures for those components in those systems that will have to be validated to some extent, okay? There's a one-time effort associated with that as well as when a PRA is updated. Those numbers will have to be updated as well.

But the rest of the data in terms of failures, unavailability, critical hours, all that stuff, is already being collected under the maintenance rule.

So from our standpoint, we do see a one-time burden associated with the setup of this program as does the staff in terms of the inspection. But longer term, once it's working, we've got one set of books, it's data that's already being collected. We're going to do the PRA updates anyway, okay? So we see over the long term an efficiency gain through this program.

Steve, you mentioned the scope of systems and MSPI. That's remaining the same with the addition of the support system element. And I think it was mentioned before that's pretty much risk significant for all the plants, okay, across the board. Component cooling water and service water are very risk significant, okay.

It has its own indicator now. I think that's overlooked sometimes in the discussion of MSPI. Rather than having to dig one level down via the cascading to what was -- what made that indicator take a hit or not, now you have its own indicator separate, again consistent with the maintenance rule, we don't cascade.

We do unavailability and unreliability monitoring on that system itself. So there is another consistency gain.

The materials issue part was raised before. We must be careful not to oversell what MSPI can do. It's not good for material issues. I think operating experience programs, the ISI program and such, are the place to address those concerns.

And I think the way that was explained this morning, that that's not intended to capture that stuff, we shouldn't oversell it and say that it has some relevance to it when it doesn't.

The system boundary definitions was mentioned this morning. We've already done this for maintenance rule implementation almost ten years ago now. That's the starting point for the scoping of MSPI. That was inspected in the baseline maintenance rule implementation program.

So we don't expect much additional effort. Probably just some -- make sure there's no significant differences and what was scoped into the maintenance rule you will capture the pertinent components.

Fault exposure time, I think that's maybe the biggest technical benefit from MSPI is getting rid of this thing as a surrogate for reliability. I mean we have a -- we already collect the actual data on reliability. Why are we still fooling around with this fault exposure term in performance indicators?

DR. ROSEN: I guess I never understood why it was a reasonable surrogate for unreliability to begin with. I mean it was -- can you explain that?

MR. PIETRANGELO: I think -- no, I can't. I'm not even going to try to offer a --

DR. ROSEN: To try and dignify it with some sort of technical --

MR. PIETRANGELO: No, no. Although as an industry, we've been using it also in the INPO and WANO indicators. So it's not like it was --

DR. ROSEN: It's been --

MR. PIETRANGELO: -- invented just for ROP. It's been around but it's time to get rid of it.

DR. ROSEN: It came into place as practice but it never had a good justification. It just kind of like tops it. It grew. And there it was.

MR. PIETRANGELO: PRA technical accuracy was mentioned. We think implementation of the program would actually do something tangible to actually improve the technical adequacy of PRAs across the board.

Unlike 5069, which is an optional rule which we hope most if not all licensees would ultimately adopt, it's going to be market driven. This thing is not market driven.

Even though it's technically a voluntary program, everybody is going to do it. Everybody is going to have to sign up for this or you face additional inspection hours.

So the actual kind of benchmarking that will be done and some of the inspection will put more attention on the PRAs and I think will improve them. And it will give more resources to the PRA because it will be hopefully part of the ROP. So I think we'll see a tangible improvement in PRA technical adequacy as a result of MSPI.

Finally, the complexity of this thing was mentioned. Is it more complex than the safety system unavailability PI? Clearly it is. Okay? But it's worth it. And it's not that complex. I think we can develop -- in fact we have even taken a first cut at the plain language description of what MSPI is.

Just like the pamphlet that was put out for the original ROP, we envision something similar here and we've already got a first cut at that. But it clearly has to be explained. I don't think it's technically complex.

We've used Fussell-Vesely measures and importance measures in the maintenance rule and in other applications. And so this is just a simple equation to put those terms together in MSPI. So technically, it is not complex.

DR. SIEBER: Well, that's sort of the in the eyes of the beholder. But I think your plain language explanation as to what the MSPI is is important because this has to be communicated to the public.

You know the ROP and it's predecessors have been relied upon by the financial community --

MR. PIETRANGELO: Right.

DR. SIEBER: -- and all kinds of other folks to differentiate one operator from another. And if they don't understand the basic framework or the structure and the meaning of these various indicators and processes, then I think that the ROP won't gain its proper respect.

MR. PIETRANGELO: I couldn't agree more.

DR. SIEBER: And I think that what you write and what the staff writes ought to coincide. I think that would be a great idea.

MR. PIETRANGELO: We've actually done some preplanning. It would require about a nine-month rollout/implementation period to do the training and communicate, get everybody on board to rollout MSPI in the industry. And the staff obviously has some resource constraints also in terms of the inspection work that would be needed to roll this out.

From our standpoint, you know, we're willing to work on whatever issues are out there be they technical, be they policy, be they resource. I think you can glean from my remarks that the industry is very positive about hopefully getting MSPI rolled into the ROP.

It's clearly a technically superior indicator than the current SSU. And it does so much in terms of bringing coherence between different processes in the regulatory framework that it's really worth it.

And with that, if there's any further questions --

DR. ROSEN: Are you going to tell us how you really -- where you really stand?

(Laughter.)

DR. SIEBER: I would point out that we have another meeting that will start shortly with the Planning and Procedures Subcommittee. And so Steve and Mario would want to leave.

On the other hand, before they do, I'd like to ask either one of them or both of them if they have any comments that they would like to pass on?

DR. BONACA: Yes, I'm impressed by the indicator and by the effort done to validate it. I think that it adds context to the information it provides that didn't exist before. And I think that context is very important.

I think that the degree to which you can represent both unreliability and unavailability and you have a weighting process based on risk significance and you have a, you know, I mean there is much more insightfulness.

I also think that there is merit to the frontstops and backstops. I mean the ability of doing some counting by the number of failures that you're looking at over a period of time tells you really the performance of the plant.

Individual event always bothered me, you know, when we were just measuring one event. And you don't know if it comes from, you know, what it really means. And, again, this adds context, it provides insightful data. I am supportive of it.

MR. PIETRANGELO: I know the staff gave

you a schedule for a potential ACRS letter on this. I do know that the committee overall has an interaction with the Commission in May also.

DR. SIEBER: That's why I wanted to know whether theirs was first or ours was first.

DR. BONACA: No, they are --

MR. PIETRANGELO: They're May 6th I believe.

DR. BONACA: Yes, that's right.

DR. SIEBER: Yes, we are first.

MR. PIETRANGELO: And given the Commission's interest in that issue, I encourage you to speak to it during the briefing.

DR. SIEBER: Thank you.

DR. BONACA: I would expect they will ask the question so --

DR. SIEBER: Yes, right. I'm going to wait for that.

DR. BONACA: You're going to be the selected person.

DR. SIEBER: Steve, would you like to add any comments?

DR. ROSEN: Yes, just a brief one. I think that the effort has been superior, the cooperation between the staff and industry is exemplary. And I think it would be a shame if given all the good work and effort that's gone on that the Commission were to back away from it at this point.

That's not to say that the concerns expressed by Bruce Boger aren't real ones. They are. And they need to be dealt with in a straightforward way and worked through if it's possible. If it's not, well then we need to know what the showstopper is but it's just -- to me that's just another part of the process of risk communication, in this case, internal.

To work through whatever the issues are and see our way clear to doing what I think we universally think is a better technical job is the way to go.

DR. SIEBER: Okay. You have to go to the meeting?

DR. ROSEN: Yes.

DR. SIEBER: We'll let you go. Why don't we go around the room and DR. Kress, do you have any comments you'd like to make?

DR. KRESS: Well, I also like this piece of technical work. I think it's good. It addresses the concerns that I and some of the other committee members have had with the current ROP process and the risk informed of the current ROP process.

It gets rid of some of the arbitrariness about the performance issues by actually tying them to what expected performance is. And it gives a chance to actually have plant-specific thresholds, which was another one of our issues.

So overall I like it and I think I'd like to see us proceed and get rid of all the problems with it and get it included in the ROP.

DR. SIEBER: Okay, thank you. Dr. Shack?

DR. SHACK: No, I'll just echo what Steve and Tom said.

DR. SIEBER: All right. Graham?

DR. LEITCH: Yes, the only thing I would add is just I like the idea of modifying the system somewhat so that as has been mentioned, that events that are recurring, repetitive events over a fairly short period of time, be factored into the process. I think that's an important insight that staff has already discussed.

And I would just like to add my support to some modification that would reflect those kinds of repetitive failures over a short period of time.

DR. SIEBER: Thank you. Peter?

DR. FORD: To use Steve's words, I'm not a PRA geek. I'm a materials geek. Within that confines, I was impressed by the MSPI program. Metrics are pragmatic. And they also meld into the existing, as I understand it, maintenance rule and tech specs.

Now my plea is that if and as this program is developed further and used, that you do have meld in or create a materials degradation capability into it. Materials degradation has been a reliability issue for many, many years. And I hope it is in the future. That's it.

DR. SIEBER: Okay. I guess my own comments are that I basically concur with my colleagues on the use of the MSPI. And -- but I do think it's complicated. I think that communication of this to the general public who is going to use it is important because the reputation of the ROP depends on the explainability of all the components that go into it.

And I think that since that's such a key document that guides the Agency's management of its compliance and enforcement actions and a document upon which licensees rely to get a measure of how they stand with respect to their license requirements, that this communication is an important factor and should be carefully considered by the staff and by the industry.

And so with that, if there are any additional comments by members or anyone in the audience or the staff?

(No response.)

DR. SIEBER: If not, I'd like to thank all the speakers, again from the staff and also from industry. And I would like to adjourn this meeting.

(Whereupon, the above-entitled meeting was concluded at 11:15 a.m.)

CERTIFICATE

This is to certify that the attached proceedings before the United States Nuclear Regulatory Commission in the matter of:

Name of Proceeding: Advisory Committee on Reactor Safeguards Reliability and Probabilistic Risk Assessment & Plant Operations Subcommittees

Docket Number: n/a

Location: Rockville, MD

were held as herein appears, and that this is the original transcript thereof for the file of the United States Nuclear Regulatory Commission taken by me and, thereafter reduced to typewriting by me or under the direction of the court reporting company, and that the transcript is a true and accurate record of the foregoing proceedings.

Rebecca Silberman
Official Reporter
Neal R. Gross & Co., Inc.

MITIGATING SYSTEMS PERFORMANCE INDEX PILOT RESULTS

PRESENTATION TO ACRS SUBCOMMITTEES ON RELIABILITY AND PRA, AND PLANT OPERATIONS

PATRICK W. BARANOWSKY (301-415-7493) (pwb@nrc.gov)
DONALD A. DUBE (301-415-5472) (dad3@nrc.gov)
OPERATING EXPERIENCE RISK ANALYSIS BRANCH
OFFICE OF NUCLEAR REGULATORY RESEARCH
U.S. NUCLEAR REGULATORY COMMISSION

April 14, 2004

Purpose and Objective of Meeting

- RES to present results of MSPI evaluation
- NRR to provide current status of MSPI
- To hear public views
- To request ACRS Letter on MSPI methodology in 2004

RES Overall Conclusions on MSPI

- The MSPI is a highly capable performance indicator that can differentiate risk significant changes in performance and addresses problems associated with the currently used performance indicators.

- The MSPI has undergone an extensive development, testing, and evaluation program. Its performance characteristics, strengths, and limitations are well understood.

- The MSPI appears to provide the best overall measure of system performance while minimizing false positive and false negative performance indications.

- The formulation is flexible and adaptable to address emerging issues and concerns regarding validity and appropriateness of outcomes.

Contents of Presentation

- Background
- Brief Overview of MSPI
- Status of Pilot Program
- Scope of Independent Verification
- Research Results on the Pilot
- Discussion of Key Issues
  - PRA Adequacy Analyses for MSPI
  - MSPI/Significance Determination Process (SDP)/Safety System Unavailability (SSU) Comparison
- Overall Assessment and Conclusions

Background

- MSPI evolved from feasibility study of Risk-Based Performance Indicators (NUREG-1753)
- NRR issued User Need Request to RES to support development of risk-informed indicator that includes unreliability and safety system unavailability
- MSPI formulated to address known issues with current indicator
  - Use of fault exposure time as surrogate for unreliability
  - Definition of unavailability inconsistent with Maintenance Rule
  - Cascading of cooling water support systems failures
  - Thresholds that do not recognize plant-specific features
- Twelve-month Pilot Program initiated September 2002
- ACRS subcommittees briefed on July 8, 2003 regarding status of pilot and RES-recommended improvements to method
  - No open items

Overview of MSPI Features

- Eliminates known problems with existing SSU Indicator.

- Accounts for both unavailability and unreliability of a system, weighted relative to their Risk-Importance.

- Uses plant PRA model to derive Risk-Importance weightings. Hence, captures plant-specific configuration and performance.

- Identifies changes in performance while limiting *false positive* and *false negative* indications.

- MSPI data are consistent with PRA methods and Maintenance Rule data. Data to be integrated with Consolidated Data Entry (CDE) Program under INPO.

MSPI Technical Approach

- MSPI monitors risk impact (i.e., change in CDF) of changes in performance of selected mitigating systems, which accounts for plant-specific design and performance data.

- MSPI consists of two elements, system unavailability and system unreliability. MSPI is the sum of changes in a simplified CDF evaluation resulting from changes in system unavailability and system unreliability relative to baseline values.

- MSPI = UAI + URI where

  UAI: system unavailability index due to changes in train unavailability

  URI: system unreliability index due to changes in component unreliability

- The risk impact of changes in mitigating system performance on plant-specific CDF is estimated using plant-specific performance data and Fussell-Vesely importance measures.

List of MSPI Monitored Systems

BWRs	PWRs
HPCI/HPCS (high pressure coolant injection/core spray)	HPSI (high pressure safety injection)
RCIC (reactor core isolation cooling)	AFW (auxiliary feedwater or equivalent)
RHR (residual heat removal)	RHR
EAC (emergency AC power)	EAC
Support System Cooling (ESW + CCW)	Support System Cooling

Outside of Scope of MSPI

- Multiple concurrent failures of components including common cause failures
- Conditions not capable of being discovered during normal surveillance tests
- Failures of non-active components.

(If any one of these conditions is present the current SDP would be used.)

Status of Pilot Program

- Twelve-month pilot completed September 2003. Preliminary draft report forwarded to NRR.
- RES performed additional analyses on PRA Adequacy (beyond original scope).
- Continued to hold numerous public meetings through the present.
- Internal NRC meetings held and position papers formulated throughout fall and winter of 2003-2004.
- RES completed draft "Report on the Mitigating Systems Performance Index (MSPI) Results for the Pilot Plants" in February 2004.
- Inspection Program Branch, Division of Inspection Program Management of NRR decides to terminate development and implementation of MSPI in March 2004.

Scope of Independent Verification

- Baseline data verified
  - Reviewed all unavailabilities, inconsistencies noted
  - Revised industry failure rates to represent 1999-2001 performance
- Current performance data verified
  - All unavailabilities compared for reasonableness
  - Reliability data compared to Equipment Performance and Information Exchange (EPIX) and Reliability and Availability Database Systems (RADS)
  - Errors noted and corrected
- All licensee Fussell-Vesely importances compared to SPAR models
- SPAR Enhancement effort completed
  - Improved eleven SPAR models to the level of component risk-importance
  - Analyzed differences between licensee PRA model and SPAR
- MSPI spreadsheet verified
- MSPI results compared using SPAR and licensee PRA
- Analyzed sensitivity of MSPI results to differences between SPAR and licensee PRA
- Analyzed MSPI, SDP, and SSU results for all component failures in pilot.

Research Results on the Pilot

- **MSPI results from pilot plant models and SPAR resolution models found to be in very good agreement**
  - **Color indication comparable if not identical depending on treatment of "frontstop" and common cause.**
  - **Numerical results within factor of three.**
- **Evaluated differences in pilot plant PRA models and SPAR**
  - **For eleven models, found three plant-specific model differences that could potentially have <u>large</u> impact on MSPI results.**
  - **Significant differences in major model inputs such as system success criteria or initiating event frequencies are primary source of significant quantitative differences.**
  - **Factors of two to three differences in basic event probability have low impact on MSPI results.**
- **Compared MSPI, SDP and SSU results for all seventy-seven component failures in the pilot**
  - **Some agreement and some disagreement, all explainable.**

Sensitivity Studies to Address PRA Adequacy

- Using SPAR Enhancement results, identified major differences between SPAR and licensee PRA model (e.g. basic event probability, initiating event frequency, system success criterion).
- Differences grouped into 3 to 7 categories of issues (e.g. AC power, AFW system, PORV success criterion, etc.).
- "Change set" generated and SPAR model re-run one issue at a time for each plant.
- Generated new PRA results including revised Birnbaum importance measures.
- New Birnbaums fed back into MSPI algorithm to generate new MSPI results.
- Quantitative and qualitative (color indication) changes in MSPI provide a measure of sensitivity of results to model differences.

Sensitivity Study Results

- **Impact of Model Differences:**
  - **Large: difference greater than 5E-7, likely to affect color performance indication given failures in the system**
  - **Medium: difference between 1E-7 and 5E-7, has the potential to affect color given sufficient failures in the system**
  - **Low: difference less than 1E-7, unlikely to affect MSPI results.**

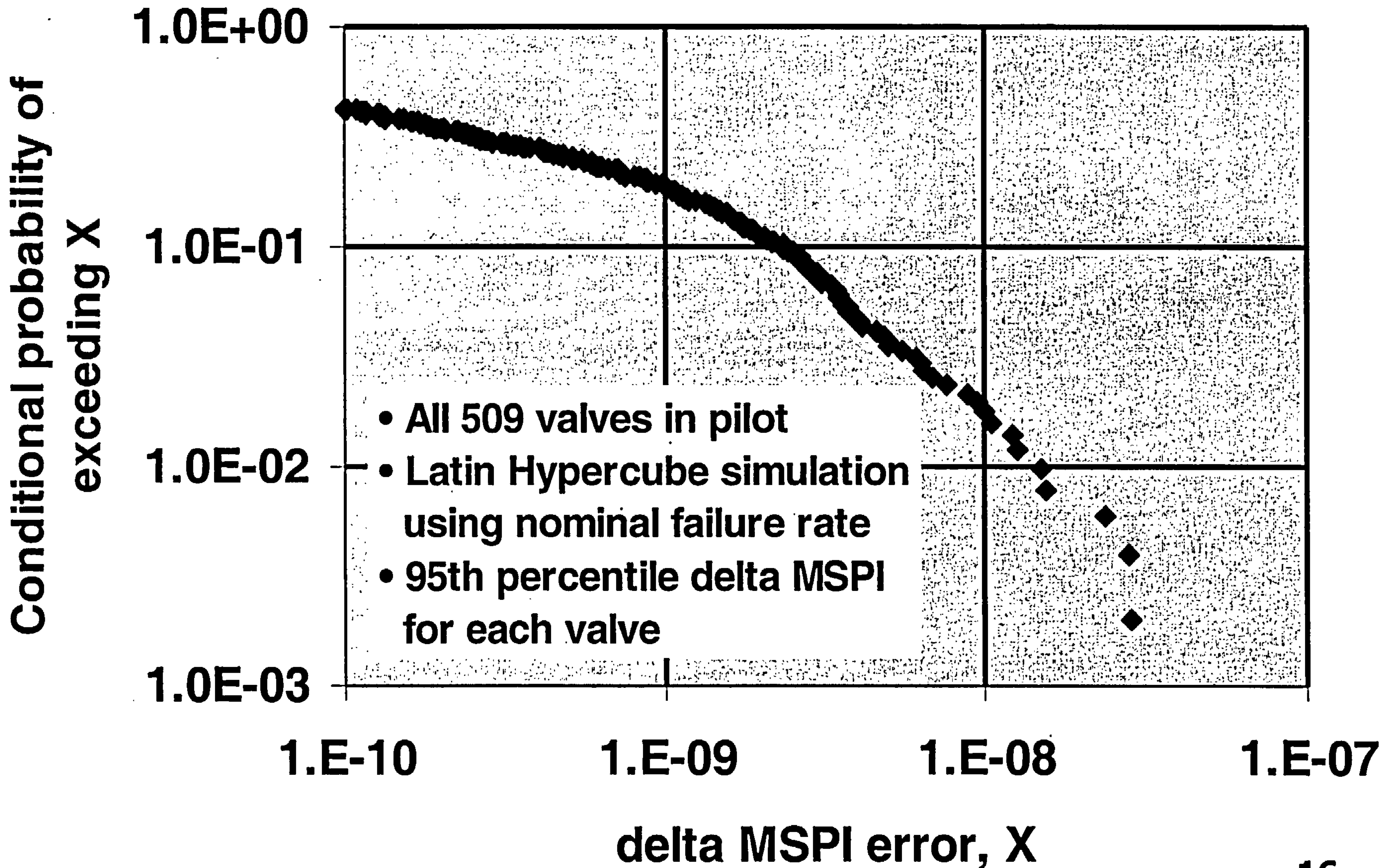
Potential Impact on MSPI	PRA Model Difference Issue									
	PORV Success Criterion	AC Power	DC Power	LOCA Issues	HPI Issues	HRS Issues	RHR Issues	SWS/CCW Issues	Power Conversion System	Misc. Issues
Large	Braidwood			Millstone 2				Salem		
Medium		Hope Creek Millstone 3 Salem	Braidwood		Hope Creek Millstone 3	Millstone 3 Palo Verde		Hope Creek Millstone 2	Braidwood Hope Creek Limerick Millstone 3	Hope Creek
Small	All others	All others	All others	All others	All others	All others	All	All others	All others	All others

(assumes all components have one failure beyond baseline)

Analysis of System Boundary Definition Error

- Guidelines require all EDGs and pumps to be included in system boundary
- Primarily an issue of omission of valves
- Probability of significant effect on MSPI:
  - A valve is inappropriately omitted, AND
  - The valve is a high risk-importance valve, AND
  - The valve has a failure rate much higher than the industry norm.
- The consequence of omission would be the underestimation of the MSPI. The valve would be subject to inspection process and an SDP evaluation of the performance deficiencies.

Analysis of Impact of Missed Valve from System Boundary on MSPI Results



MSPI/SDP/SSU Comparison

- They are fundamentally different approaches:
  - MSPI measures statistically valid risk-informed change in performance of systems over 3-year rolling intervals.
  - SSU directly accounts for unavailability, and indirectly attempts to address unreliability through use of fault exposure time, also over a 3-year rolling interval.
  - SDP measures short-term risk significance of a failure or condition associated with performance deficiency.
- Compared results for all 77 component failures in pilot.
- All non-Green SSU driven by fault exposure hours, one because of T/2 assumption prior to change in ROP Guideline.
- SDP non-Green findings for single failure often driven by short assessment period (less than a year) with insufficient data to measure statistically valid change in performance.
- MSPI White or near-White indicators usually involved multiple failures and unavailability providing high degree of confidence of adverse change in system performance.

MSPI/SDP/SSU Comparison (cont.)

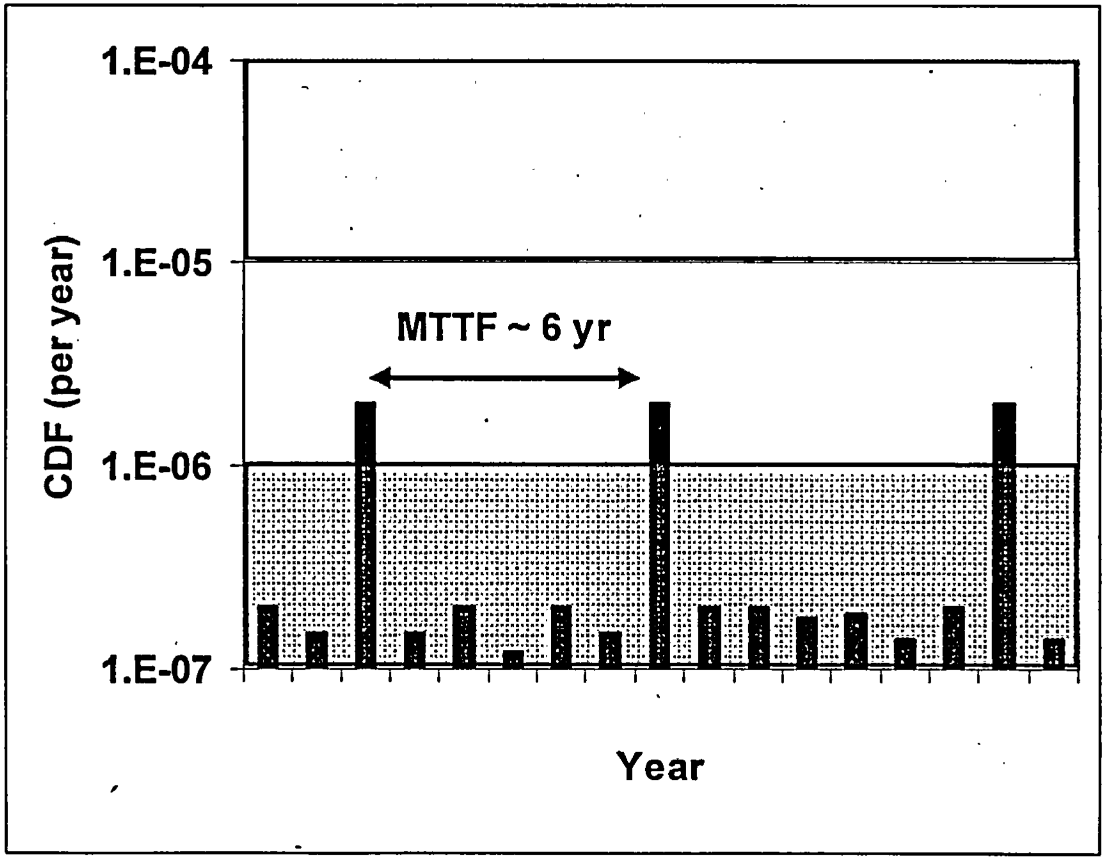
- MSPI captures as many if not more reliability/availability performance degradations than SSU/SDP
  - Historical review of 1,659 SDP findings and 5,157 SSU quarterly indicators.
  - Only 0.5% of SSU indicators have been non-Green in 3.25 years reviewed.
  - Average of 4 non-Green SDP findings per year for mitigating systems related to actual single failure (not degraded or non-conforming condition) for all 100+ nuclear units.
  - MSPI pilot resulted in 2 White indicators, and 3 near-White (data collection stopped 1st Qtr 2003)
  - SDP had 2 White findings, where MSPI was Green or near-White.
  - Analysis and numerical simulation (Latin hypercube) shows equal to or greater non-Green indicators with MSPI.

Most Important Comparisons of MSPI, SDP, and SSU

Plant	Component	MSPI	SDP	SSU Value/Threshold	Comments
Braidwood-1	3 failures of AFW diesel pump	2E-06	One Green finding	2.5% / 2.0%	MSPI White comes from combination of unreliability and unavailability.
Hope Creek	3 failures of HPCI MOVs	1E-06	None	1.7% / 4.0%	MSPI White comes from combination of unreliability and unavailability.
Palo Verde-2	1 failure of motor-driven AFW pump	4E-07	None	0.5% / 2.0%	*Frontstop* applied. One more failure results in White MSPI.
San Onofre-2	6 failures of salt water pumps	*backstop* of 7 makes this a near-White	None	N/A (support system)	Calculated MSPI near zero owing to balancing of unreliability and unavailability.
Salem-1	4 failures of EDGs in 3rd quarter 2002	8E-07	9E-06	1.5% / 2.5%	One additional failure through 2nd quarter 2005 or net addition of 40 hours of EDG unavailability results in White MSPI. RES is assessing adding short-term *backstop*.
Millstone-2	1 failure of turbine-driven AFW pump	-4E-07	3E-06	2.7% / 2.0%	Motor-driven pumps much better than baseline for UA and UR. Turbine-driven pump at baseline for unavailability. SDP & SSU results driven by fault exposure time.

Application of Fault Exposure Time to Single Component Failure Can Lead to Volatile Indication of Performance Otherwise at Industry Norm (i.e. *False Positive)*

Example: Typical turbine-driven AFW pump tested monthly with assumed 14-day fault exposure time
Prob(FTS) = 1E-2
MTTF ~ 6 years (includes demands from additional operations)



MSPI Qualities

- **MSPI addresses both false positive and false negative concerns**
  - **Constrained non-informative prior (CNIP) for component reliability, *frontstop,* and *backstop* appropriately constrain minimum and maximum failures to White.**
    - **CNIP demonstrated to provide best false positive/false negative characteristics of priors considered in NUREG-1753.**
    - **Frontstop minimizes likelihood that one failure beyond baseline in 3-year period results in White. However, index could still become White with one or even zero failures if there is significant system unavailability.**
    - **Backstop results in White indication if component type exhibits statistically significant departure from nominal industry failure rate, regardless of risk-significance.**
  - **Latent faults and multiple concurrent failures with high risk would continue to be evaluated by SDP.**

MSPI Qualities (cont.)

- MSPI is consistent with Maintenance Rule (MR), Tech Specs, and principles from SECY 99-007, "Recommendations for ROP Improvements"
  - Consistent with MR definition of failure, demand, and unavailable hours.
  - Consistent with the basis of Technical Specifications and the Maintenance Rule in that single failures that are detectable by normal surveillance are neither risk-significant nor indicative of performance degradation.
  - "Sufficient margin...to allow opportunity [for licensee] to take appropriate action."
  - PIs should be used where practical, and inspections cover all other areas.

In Conclusion

- MSPI has been thoroughly tested, evaluated, and discussed in numerous public meetings.
  - It addresses problems with currently used PIs.
  - Its capabilities, strengths, and limitations are well understood and accounted for.
- With a few peripheral adjustments, MSPI development and evaluation will be complete.
- Even in current configuration, MSPI is a highly capable performance indicator.
- MSPI has highly desirable qualities with respect to:
  - Plant-specific risk implications
  - Proper treatment of reliability and availability
  - Ability to capture performance degradation and address false positive/false negative concerns. Provides robust results.
  - Computation is structured and programmable.
  - MSPI is consistent with Maintenance Rule, Technical Specifications, and ROP SECY 99-007.
- PRA adequacy issues are identifiable and manageable; few contribute to significant discrepancies in outcome.

Current Plans to Complete MSPI

▪ Internal comments due	May 2004
▪ Public comments due	June 2004
▪ Prepare draft final MSPI report	August 2004
▪ ACRS full committee	September 2004
▪ ACRS letter	TBD
▪ Issue Final MSPI Report	October 2004