Official Transcript of Proceedings ACNWT-0166

NUCLEAR REGULATORY COMMISSION

ORIGINAL

Title: Advisory Committee on Nuclear Waste 146th Meeting

PROCESS USING ADAMS
TEMPLATE: ACRS/ACNW-005

Docket Number: (not applicable)

Location: Rockville, Maryland

Date: Tuesday, October 21, 2003

Work Order No.: NRC-1140

Pages 1-63

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

TR08

UNITED STATES OF AMERICA

NUCLEAR REGULATORY COMMISSION

+ + + + +

ADVISORY COMMITTEE ON NUCLEAR WASTE (ACNW)

146TH MEETING

+ + + + +

TUESDAY,

OCTOBER 21, 2003

+ + + + +

ROCKVILLE, MARYLAND

+ + + + +

The meeting convened in Conference Room T-2B3 of the Nuclear Regulatory Commission, 2 White Flint North, 11545 Rockville Pike, Rockville, Maryland, at 10:30 a.m., B. John Garrick, Chairman, presiding.

MEMBERS PRESENT:

B. JOHN GARRICK   Chairman, ACNW
MICHAEL T. RYAN   Vice Chairman, ACNW
GEORGE M. HORNBERGER   ACNW
RUTH F. WEINER   ACNW

ACNW STAFF PRESENT:

JOHN T. LARKINS	Executive Director, ACRS/ACNW, Designated Federal Official
SHER BAHADUR	Associate Director, ACRS/ACNW
NEIL M. COLEMAN	ACNW
HOWARD J. LARSON	Special Assistant, ACRS/ACNW
MICHAEL LEE	ACNW
RICHARD K. MAJOR	ACNW

ALSO PRESENT:

TINA GHOSH	ACNW Summer Intern
ANDY CAMPBELL	DWM
PHILIP JUSTUS	DWM
TIM MCCARTIN	DWM
TIM KOBETZ	NMSS/DWM
W.M. BURTON	NMSS/DWM
CHRIS GROSSMA	NMSS/DWM
KEITH COMPTON	NMSS/DWM
E.V. TIESENHAUSEN	CCCP
NORMAN HENDERSON	Bechtel SAIC Company
CAROL HANLON	DOE
LEM REITER	NWTRB

I-N-D-E-X

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

10:33 a.m.

CHAIRMAN GARRICK: Good morning. The meeting will come to order. This is the first day of the 146th meeting of the Advisory Committee on Nuclear Waste.

My name is John Garrick, Chairman of the ACNW. The other members of the Committee present are Mike Ryan, Vice Chair, George Hornberger and Ruth Weiner. Somebody is missing today, as you can observe, and it is with mixed emotions that we note that Milt Levenson submitted his resignation from the Committee effective October 10, and we wish Milt the best in his future endeavors and thank him for his many efforts for the Committee.

During today's meeting, the Committee will hear the summer intern's final report on assessment model uncertainty and performance assessment. We will review plans for the Biosphere Working Group, finalize proposed activities for the November 18, 2003 trip to Yucca Mountain and Amargosa Valley. And during this afternoon's retreat session continue identifying topics we intend to examine over the next 12 to 18 months. As noted in the agenda, some portions of that session may be closed.

John Larkins is the Designated Federal Official for today's initial session. The meeting is being conducted in accordance with the provisions of the Federal Advisory Committee Act. The Committee has received no requests for time to make oral statements from members of the public regarding today's sessions, and should anyone wish to address the Committee, please make your wishes known to a member of the staff.

It is requested that speakers use one of the microphones, identify themselves and speak with clarity so that they can be heard.

Before starting the first session I'd like to cover some brief points of interest. Jenny Gallo, Chief of the Operations Support Branch, ACRS/ACNW, has been selected to participate in the NRC Leadership Potential Program. This is an honor. One hundred seventy-five employees and she was one of the 25 selected.

Recent Agency announcements of interest to the Committee include Dr. Keith McConnell has been appointed Executive Assistant for Materials and Security in the Office of the Chairman. He served in increasingly responsible positions in the staff and is currently completing the requirements of the Senior

Executive Service Candidate Development Program. He has also served on the personal staff of former Chairman Ivan Selin and former Chairman Richard Meserve. Keith has appeared before the ACNW many times. James Dyer has been appointed Director, Office of Nuclear Regulation, and James Caldwell, Regional Administrator for Region 3.

The appropriations bill for DOE and the NRC has not been signed into law. The Agency is funded until October 31 of this year under a contingency resolution. And, finally, the Trade Press has indicated an agreement between Congress and the White House to nominate both Gregory Jakskul, Senator Reid's chief license advisor, and retired Vice Admiral John Rosenbacher to the NRC, filling the seats left vacant by the departure of Richard Meserve and Greta Dicus.

If there are no further questions or comments, I think we'll proceed right on with our agenda, and, Tina, we'd like to hear how your project has taken shape.

MS. GHOSH: Is it fine if I sit? Everyone can hear me?

MR. LARKINS: Or you could stand and record it.

MS. GHOSH: You want me to?

MR. LARKINS: No, no, no.

(Laughter.)

MS. GHOSH: Thank you so much for allowing me to come back and speak with you. I really appreciate the opportunity. And while this is my final report to the Committee for what I did over the summer, most of you know my thesis is still going on and I expect to finish in May. So whatever comments you have I would incorporate into my thesis, and you'll get a more final product in a few months. So thanks again for allowing me to speak here today.

So I think everybody remembers what my topic is. The main points were probing the effects of the model uncertainties in the Yucca Mountain performance assessment, and we talked about taking a scenario-based approach to first identifying what things might be important and then assessing better how likely they are to occur and the consequences and prioritizing, for example, research for further studies, uncertainty studies, based on that.

So the first thing is basically to find out what is actually causing the risk in the system. So I've been using an older version of the NRC's TPA code, so I'll just show you, the first thing I started with was a base case run from the TPA code. And this is much more conservative than what the current databases show, because there were -- as you know, the code has become more realistic over time. This is based on the TPA 4.1 code. And once again, this is just a base case, so it doesn't consider igneous activity or human intrusion. The seismic mode was turned on so there may be effects from rock fall, but there's no vulcanism just so I could focus the study.

So we see here that in most cases we do not exceed the regulatory limit of 15 millirems per year, and I just the Neptunium 237 dose for simplicity, so we're ignoring all the other radionuclides. And if you actually look at the doses in the 10,000 time frame, I mean it's typically the Neptunium 237 which is the dominant one. So on the bottom we have the time scale from 4,000 to 10,000 years, which is 10,000 years, the regulatory compliance period, and for the most part the doses are well below the limit.

The green curve shows you the mean which comes in around two millirems, but we see that there are just a few realizations that actually exceed the regulatory limit of 15 millirems. In this case, there were nine cases where we found that the dose was exceeded. In about 80 cases, there was no dose at all for 10,000 years, so those curves are not shown on the graph. And as I said, for the vast majority of the cases, we're well below the regulatory limits. Since this is a log graph, all the gray, the dense gray lines below the limit are actually quite a bit below the limit.

So the point is now that we see that there are just a handful of scenarios that might exceed the dose, it makes sense to focus on just those to see what might actually cause the risk in the system. And by risk I guess how I'm defining it is that we can actually exceed the consequence of 15 millirems per year, because we don't really care about all the case where we don't even come close to that.

So this is just a very crude sort of first cut at building the scenarios that might cause and exceedance of 15 millirems, and what I focused on were the factors that were basically identified by NMSS, in particular Tim McCartin's tracing studies, for those attributes of the repository which might allow a dose of 15 millirems to occur. And just to get a very crude idea of whether we can explain those realizations that we found using just those factors, I tried to look at where in the parameter distribution range the realizations came out.

So if we look at the case with the highest dose, which was about 100 millirem per year, we see that the waste package flow factor, which sort of controls how much water that is coming in from infiltration actually gets to the waste package, and the waste package defect rate, which is early defects, not from general corrosion but some other early waste package failure, and the neptunium solubility are all assumed to be very high. I mean we're at the 99th percentile, 98th and 97th for these factors.

In addition, the sorption coefficient in the unsaturated zone for the Calico Hills unit is at the 64th percentile. Did you have a question -- oh, sorry. Where you see blank spots is basically where the parameter value doesn't seem to explain anything because it's below the median. So just to get a very crude estimate now of the probability, I just multiplied the exceedance probabilities of these factors if they were to explain the dose to see what kind of exceedance probabilities we might have.

Now, the first thing is because I haven't done a lot of runs, this is not a very good basis for saying that these are the scenarios that might actually exceed it, but this is just what I've done so far, and I'll talk a little bit later about the other sensitivity studies I've done to sort of capture the classes of scenarios that we might actually find that make us exceed the dose limit.

But just using these factors I think some of the runs we can't explain by these attributes, so one example is the run where we have 20 millirem. It doesn't seem that there's enough there to have us actually exceed the dose. There's something else going on there that isn't captured by these attributes. And one notable thing that I left out was looking at the waste form dissolution rate which actually is definitely found to be important and also the well pumping rate for the dose receivers which I've also left out. So it's possible that those provide some explanation.

The other reason these are very conservative is I think the drip shield was turned off, so in fact we don't have the drip shield in this particular set of runs. But, basically, I'm just trying to develop the methodology for identifying scenarios and then seeing how we can evaluate it. So that's why -- I mean they're more conservative than would be.

So the thing is now if we want to look at what risk we're actually getting from these scenarios, we can plot a risk curve in terms of the complementary, cumulative distribution function for exceeding particular doses, and if we define the universe of possibilities to be just from this 200 base case run, which is not but just as a first cut, we found that nine of the realizations exceed the dose, and I just plotted the actual doses that you get from these runs.

CHAIRMAN GARRICK: Tina, what's the total number of realizations?

MS. GHOSH: It was 200.

CHAIRMAN GARRICK: Two hundred.

MS. GHOSH: Yes. So each realization, you know, approximately a half percentile.

Now, the reason that I bring this us I realize that the licensing criterion is in terms of the peak mean dose for any given year, but the thing is we might -- we want to look at the full risk spectrum for other reasons, for example, for prioritizing future research that we might require, might want to do to get a better understanding of what risk can come out of the repository.

So the next thing I want to propose is that maybe we can consider some additional risk goals for these purposes, not for necessarily for the licensing criterion but just to give us an idea of how comfortable we are with what we project for the repository. Now, the first goal that I'm showing here is an adaptation of what was used for WIPP in the EPA rule which was basically -- and I realize the WIPP consequence is total different because they looked at cumulative release over the 10,000 years, so you just had one number rather than calculating a dose for every year. But they said, "We want to make sure that we don't exceed the goal, that there's less than ten percent of exceeding this goal and that there's less than, I think it was, one in 1,000 of exceeding ten times the goal."

So all I've done is I've graphed what the equivalent would be for Yucca Mountain, and of course our risk curve here is something totally different, because basically you can think of the curve as a slice in time, and the CCDF for that slice in time actually is peak dose, and in a couple of cases the peak dose occurs before 10,000 years but that's not quite right. So it's not quite right but it's approximately.

And for the WIPP, of course, you actually had a whole family of CCDFs and they were comparing this for the mean. But, any rate, I think the general idea is there. So this is one possible goal. This isn't defined right now, so it's something that would have to be defined.

The other thing we could do is think of it more like a farmer's curve, which would be something closer to this line, and the slope of the line I think we could build in terms of societal preferences for risk. I think the Dutch government has done something like this. And industry -- if you look at industry's sort of revealed preferences, typically the type of risk averse curve like the diagonal one in the slope, I don't have good numbers right now for this slope but it typically tends to be about minus 1.2, which basically means that for an order of increase in the consequence you want more than an order of magnitude decrease in the probability of achieving that. So this is just some ideas to keep in mind, and I want to talk about what we want to do in terms of this idea.

So if we look at now how -- what do we want to do? We want to compare our assessed risk so far with the performance assessment in terms of the risk goal. And if we are very far below this risk goal, we can be pretty comfortable if we're confident in our estimates. And as we get further away from the risk goal, of course we're increasingly more uncomfortable. So this is just the basic idea.

Now, in this case, we see that the CCDF is very far below our risk goal. I mean it's very far below. So the point is why would we still worry about it if we're so far below? The thing is we know that we've left a lot of things out of the analysis for simplicity or for different reasons, and we discussed this before but, for example, there might be alternate models available that we haven't incorporated into the PA, their dependencies are coupling, which are omitted, there are sometimes inconsistencies in the way that the sampling is done and propagated, there's uncertainty introduced in the model abstraction process, and a lot of our assessment is based on expert judgment or experts' interpretation of evidence from lab studies and so on, and there may be biases and overconfidence in these interpretations.

And the worst of all is the incompleteness, is that even if we feel good about everything else, frankly, we still always just don't know what we don't know, and it could be a troubling thing. And I really want to motivate the incompleteness part because one thing I hear so often is that our assessed risk is so low below the criterion let's just stop worrying about it all together.

And so why do we worry? And I just wanted to bring up a couple of examples. This one is very dramatic, but you probably remember the Mt. St. Helen's eruptions from 1980. Before 1980, there -- basically, there were a lot of volcanologists studying the Mountain, because they knew it was active, they knew it was going to erupt pretty soon, but the thing is that -- so you have the picture before and after. And it was just amazing that even with all these people studying it, the actual eruptions stunned the entire community. They had no idea about the magnitude of the eruption, they didn't know the direction it would explode, they always thought volcanos explode up but in fact this one blew out the side of the Mountain, and the actual consequences of the eruption stunned the entire scientific community. And David Johnston, a USGS volcanologist, actually died because his monitoring station was too close to the eruption. And the reason I bring this up is that obviously nobody would just sacrifice their life. They truly believed that it would be much smaller, and it was a shocking surprise.

And on a more mundane level, I've looked at some of the expert judgment cases where we have some assessments afterwards of how well experts have predicted various variables, and this one is taken from a benchmarking study that the European community did. And what they did is they simulated a severe nuclear reactor accident. Basically, it was a simulation of a partial core melt accident, and they asked some of the top experts in the field to predict some key variables, like the time to the peak pressure, what the peak pressure would be and so on. And a surprise in this case is identified as the actual variable being outside of the 90 percent confidence interval.

So it's either below the fifth percentile or above the 95th percentile. And so one would expect that there should be ten percent of surprises, but in actuality if you look here, the turquoise bars are the actual assessments, the number of assessments, and the red bars are the number of surprises in the assessments. And, on average, there is about one-third surprise rather than ten percent, as one might expect. The aggregate was much better, it was actually ten percent, but it's still some cause for concern, because depending on who you asked about the variables and sort of assessments for the variables.

CHAIRMAN GARRICK: Isn't the key to that, though, what is the evidence that each of the experts are looking at to form their opinion?

MS. GHOSH: Yes. You're absolutely right. And there was a lot of setup that went into the elicitations. So, for example, they're very precise about past experience. The experts were allowed to do whatever calculations that they wanted in order to come up with the estimate. Because, yes, that's always a concern, but there was opportunity for the experts to make a lot of the analyses and they shared the information base. Is that your question?

CHAIRMAN GARRICK: Yes. I'm thinking of Ed James' definition of probability where he says something to the effect that probability is subjective to the extent that it's not a property of the real world, but it is objective in the sense that if we're all given the same information, we're all basically wired the same, we will surely assign the same probabilities.

MS. GHOSH: Right.

CHAIRMAN GARRICK: And so the conversion from subjectiveness to objectiveness comes about not by just taking the word of the expert but by examining the basis of the expert's opinion and finding when you get to that level that most of the variation between the experts comes about as a result of a different knowledge base that each one of them is considering.

MS. GHOSH: Right. I think -- thank you for bringing that up. I think this is one of the key issues in how expert information is used in general for risk assessments. And a lot of the elicitations they try as much as possible to pool all of the knowledge base from the experts so that everyone shares the same information, but the truth is it's very difficult to do that. And one of the things that I've been interested in is trying to figure out which piece of the disagreement is due to different information that experts possess versus different interpretation of the same database. And I haven't -- I really want to pick that apart at some point, but I haven't gotten that far yet. But it's true that even -- you can find people who are maybe trained at the same institution, that have the same information base and give them a set of experimental results and they might interpret it differently because they think whatever, for whatever reason. But I think it's a very interesting issue, and I would like to work on that at some point. But for this, I wasn't part of this study so I can't tell you how much information they shared, but at least on the surface they were supposed to have shared the same information base.

MS. WEINER: Could I ask -- I have a question for you.

MS. GHOSH: Yes.

MS. WEINER: Do you know how this success rate or failure rate, or whatever you want to call it, how this compares to other expert assessments? Because just my basic recollection of what I know about expert elicitation is that even 33 percent isn't bad.

MS. GHOSH: Yes.

MS. WEINER: Ten percent is really pretty good.

MS. GHOSH: Right. You're absolutely right. So from what I understand, in a typical assessment, if you ask people to give you the fifth to the 95th percentile, often what you end up getting is the 25th to the 75th percentile. But the thing is if you look at how the distributions are actually used, for example, I've tried to figure out how the DOE uses assessments in its PA. They don't adjust for that, they just take the distribution as it is. And in terms of this type of study, because a lot of the expert elicitations are for situations we cannot actually produce in reality, like even the risk simulated a reactor accident. We've never had like a real reactor accident, but there have been cases where they've asked experts to predict the distributions for what the analysts call seed variables and then look at the performance of the expert assessments on those seed variables.

And a lot of this was done was a joint NRC study with the European community where Delph University basically developed a method of sort of calibrating expert judgments based on their performance on the seed variables. So there is some data out there in terms of how experts perform for predicting different variables, but I don't have all the data right now. I want to get it very soon but I don't have it yet. And this data was basically for like cesium transport studies because they were looking at potential severe accidents from reactors. So there's some key radionuclide transport variables that they were tested on. So I don't have the data yet, but I want to get it.

Okay. So one of the things that I want to talk about is basically in order to prioritize research, which is maybe one of the main risk-informed activities, what are the importance dimensions that we want to look at? So, obviously, we want to look at the assessed risk, but we don't know how confident we are in those risk estimates so we want to develop some systematic ways to assess our confidence in the estimates and the likely nature of any biases that might be in the analysis. We want to consider incompleteness, which might be part of our confidence thing but I define it slightly differently. And also we want to see how we're doing in terms of defense-in-depth, which one could think of as a strategy to deal with incompleteness, but we may want to look at performance on independent measures for defense-in-depth as well. And the last thing I just threw in, I don't want to talk a lot about today, but we should consider the public concerns in confidence just because if you look historically at what's driven decisions in the program, public concerns have often driven some key decisions, and there might be a better way to sort of anticipate the concerns and assess the likely outcome of different choices if we think about it systematically.

So the first thing I want to do is think about how we can build confidence in the estimates that we have. So we have -- I have this very crude CCDF curve for the few runs that I did, and what can

I say about where the biases might lie in terms of that curve? Now, in this case, I'm talking in generalities because I haven't probed all the information basis for the specific attributes yet, but these are just some, I think, obvious things. So the first thing is that the initial defective waste package rate is postulated to be very high in the TPA 4.1 database if you compare it to what the DOE is that I'm talking about. So, for example, if you have a 0.97 percent initial defective rate, that gives you something on the order of 100 early waste package failures in the 10,000 year period, which if you look at the DOE assessment, basically the probability is zero of that happening. So the waste package failure rate is one place to probe.

And from what I understand in the TPA database, the waste package failure is sort of a binary thing. And once it fails the inventory is available to whatever water there is to carry radionuclides away. So if we look at that compared to what the DOE assumes, the DOE assumes about 90 protection of the waste form based on the cladding. And although -- I just wanted to point out that one review panel has said that maybe that 90 percent is a little bit optimistic because there is some effects that are left out of the analysis, such as the basket corrosion that might affect the cladding, but at least maybe that gives us an order of magnitude maybe.

And with the bathtub emersion model, that's probably also kind of a conservative effect and also a place where there's an inconsistency, because the bathtub is assumed even when there's not enough water to actually make a bathtub effect.

Another example is I checked in the EPA rule what the representative volume was supposed to be for the receptor group and looked at the average dilution volumes that were calculated by the TPA runs that I did, and I think part of the discrepancy comes from the fact that I looked at the ten kilometer point rather than the 20 kilometer point, because I don't know where the point is supposed to be at this point, because the EPA rule is in terms of a very specific latitude and I don't know how far that is from the repository but the DOE had been using 18 kilometers. So maybe that's the reason for that discrepancy. But it's about five percent of what the EPA's representative volume is, so of course there's a lot less water to dilute the concentration of radionuclides. So just looking at these couple of things maybe there's a potential reduction of the dose by about two orders of magnitude.

So if we look at potential optimistic assumptions, there's assumed to be no general corrosion failure before 40,000 years in the TPA, and I think it's even further out in the DOE model. And there is a good basis for that in terms of industry experience with the materials and so on, but once again there might be some things that we just haven't figured out yet because of incompleteness that may affect how we assess that time frame.

The other thing is something I hear people talking about a lot on and off is that whether there are possible groundwater fast paths that we just haven't found yet, and, actually, I say faster paths because some accelerated travel is already simulated in both the DOE and the NRC's performance assessments. So in order to assess this, I mean what's one of the ways that we can estimate the order of magnitude? We can look at all the past cases where we've been surprised by the fast paths, and there are actually a lot of data out there from historical cases, from low level waste sites, from most of the DOE complex sites where there is some contamination traveling maybe from Chernobyl, and the natural analogs, which is already considered quite explicitly in the PAs.

And I don't have the numbers yet, but just to give you an example, if we say maybe about a thousandth of the nuclides can travel about ten times faster and we see how much larger the dose could be, we can just take a weighted fraction of what's coming out of the unsaturated zone and a very crude estimate is that maybe there's about a three percent increase in the dose which is not very significant.

So one of the issues that might be of concern, and actually we just talked about this a little bit, is that when experts sort of disagree on how to interpret the same information, because that's a potential cause for concern because it illustrates that there isn't a consensus about what's actually going on. And if we consider the DOE and the NRC as two expert communities, we can see that there is some disagreement about the relative capability of the unsaturated zone versus the saturated zone to attenuate the radionuclide plume, and so the question is why? Is it because the DOE has made some conservative assumptions just for simplicity because they don't have the resources to study everything? That would be less cause for concern. Or is there really a legitimate disagreement about what's going on in the unsaturated zone and the saturated zone and what the dominant processes are for release, which would be more of a concern.

And the next thing I want to talk about is there's a lot decoupling that goes on in terms of the evaluating the capability of different barriers in the system and different attributes. And I just want to -- well, eventually, I'd like to find some systematic way to see what it's okay to decouple versus what it's not so okay decouple. And the first place that I looked is basically the NMSS' latest risk baseline document, or the draft that I have from August, in terms of the key technical issue agreements and how they affect the release from the engineer barrier system, the unsaturated zone and the saturated zone. And all I did was just try to map the issues to how they affected different attributes which eventually affect, for example, the release from the EBS. And I apologize for this slide; the justification is kind of messed up. I hope you can read it on the copy that you have in front of you. So the main thing is if you look at sort of the issues that are affecting the release from the engineered barrier system, a lot of it is connected, there are a lot of feedbacks, and so there's a lot of potential for synergistic effects.

If we look at what's affecting the

unsaturated zone and the saturated zone, which are the next two slides, it's very difficult to make a case that when we look at the scenario as a whole that somehow the capabilities of the unsaturated zone and the saturated zone could affect what's coming out of the engineered barrier system. So I think a tentative conclusion is that probably for sure we can consider the unsaturated zone and the saturated zone capability independently of the engineered barrier system, but because what's happening -- everything that's happening up to the neptunium release from the engineered barrier system is somehow connected.

For example, how much water is coming in and where it travels down affects the chemistry, which affects the corrosion rates. If you look at coupled effects of the temperature operating mode, you have some cyclic effects. So it's kind of difficult to evaluate things separated from each other. And what I would say is that it's difficult to evaluate the barrier capabilities without looking at the whole picture together.

So in our confidence building studies, we may want to consider these coupled effects in reevaluating the barrier capabilities. And why do I bring this up? When I looked at the supplementary studies, for example, for barrier capabilities, typically what you see is the flux going into a particular barrier versus coming out of the particular barrier, but these are all based on the assumptions in the performance assessment, so there might be a lot that's left out of the analysis, and I think it could be worthwhile to probe some of these.

So in terms of defense-in-depth, which I think people are pretty concerned about, I wanted to define it maybe a little bit more concretely than what's discussed right now. So the first thing that I would like to do, and I'm sort of basing this on the Reg. Guide 1174 framework, is if we look at the risk curve picture versus the risk goal that doesn't exist right now but that we can postulate, we can think about, for example, how much safety margin we have depending on how far we are from our risk goal. If we're very far into the comfortable region and we're confident in the assessments, we can be more comfortable about the repository.

The other things that we might have as tests are something that's similar to the single failure criterion of reactors which is is there any single assumption that if it's wrong could defeat the system? And if you look at sort of the scenarios based on the realizations I had done, there was one case where the only explanation factor seemed to be a very high percentile for the sorption rate in the unsaturated zone in Calico Hills unit. If it's true that just having a very low sorption rate for that unit is enough to maybe create a dose, that would be cause for concern. I'm not sure if that's the case yet because I don't know what else is going on in that realization I need to figure out, but that's something to consider.

MR. HORNBERGER: But it's just one realization, so you already know that that isn't the case, that is a low KD in the Calico Hills --

MS. GHOSH: Yes.

MR. HORNBERGER: -- could occur in thousands -- not in your case, but lots of other realizations and you don't get a dose. It's certainly not a factor that would cause failure.

MS. GHOSH: Yes, it shouldn't, but the thing is in that particular case that was the realization at the highest range for what we stratified. So, yes, I'd have to go back and test that. Actually, I tried to do some of that. Maybe I should just tell you that right now. What I wanted to do eventually is to find classes of scenarios that would fail our system, because all I've done is the base case and it's not very reliable. But the reason I didn't show those results is that it's very difficult to find, I found -- I could have guessed this from the beginning, but it's very difficult to find sets of things that give you failure with very high probability.

And even when I forced the stratification into very high ranges for a couple of factors, I still always got zero dose in 40 percent of the cases, and I have to figure out why that is. And still in most cases you didn't exceed the limit. The most I got was a 30 percent chance of exceeding the limit, but of course the doses were much higher because where I had pushed the sampling. So I don't think it's the case that you could fail the system with just that, but it's just something to consider because that was the realization that had the fewest elements that seemed to be in the range, but I don't know what else is going on. I would like to figure out very soon.

And I guess in terms of a diversity goal for defense-in-depth along similar lines, we could look at the scenarios that have the fewest elements and processes and see how confident we are in the probabilities assessed for those. And I hope that makes sense because if we're relying on too few elements, well, I think it's obvious.

So in terms of qualitative aspects -- so for those assumptions now that we find that are important, what are the technical basis for these assumptions? We can feel -- well, we can assess the quality of these by some qualitative factors. So, for example, are they based on laboratory experiments, how much inference is involved in the assessment, and I think the waste package corrosion studies is a good candidate to investigate these just because I've seen a lot of commentary that although there have been extensive experiments done, perhaps not for the full range of conditions that might exist in the repository, so there's a lot extrapolation involved rather than getting evidence for all of the ranges of conditions.

The other thing that I think people care about, and especially from a public confidence point, is what kind of peer review was used and who was involved? And one thing that's been in the papers recently is the environment -- at the government bodies that are conducting these analyses, is the environment conducive to people raising safety concerns? And I just bring this up because it's actually been in the papers over the summer and there have been some issues about people reporting that they don't feel comfortable raising safety concerns, which really doesn't do a lot for building public confidence.

And the other thing I wanted to bring up is that there's been ongoing public debate in the papers about Yucca Mountain and the topic sort of has -- topics have evolved over time. I think the water one is probably still around but it was one of the early ones. There's a lot of questions about where does it come from and where does it go. The igneous activity is probably an ongoing one. I think the criticality issue was put to bed and I'd be very interested to know how that was done, because that was an issue that was in the papers a lot for a while but not anymore, and I think recently there's a lot of questions about the corrosion rates and mechanisms for a waste package and the drip shield because there is a lot of emphasis on those barriers for the performance of the repository.

So the thing is, is it worthwhile to take these in to consideration when we prioritize what we're going to do for further studies, and I think it is because if you don't, you end up having to make reactive decisions later that might cost more or you might have experiments where you don't actually achieve public confidence with what you've done.

And I think that when you look to further studies and prioritizing what we should do, it's worth it to evaluate sort of how the experiments are going to affect our assessment along all of the importance dimensions. And the thing is we're never going to get perfect information from whatever experiment that we do, so it's worthwhile to try to evaluate what we're going to get from the imperfect information and get some kind of distribution of how we're going to do on these various attributes like defence-in-depth and public confidence when we assess what we should study.

I think from based on what I heard over the summer the DOE is definitely developing this, but the focus and the methods are slightly different. So it might be worthwhile to reevaluate that.

So as I said, what do I eventually want to do? I think all of this was based on a very small sample of what might actually happen with the repository. And eventually I would like to come up with these classes of scenarios that, at the end state, we could say we're going to exceed the dose -- we'll be in this dose range with this probability, because I think we're never going to have a set of scenarios that we can say deterministically will definitely give us this dose, just because it's not worthwhile to look at all 300 parameters to define those.

So I'd like to bin them according to what our preferences are for risk. So not just looking at the 15 millirem cutoff, but maybe looking at which things cause us to be 15 to 150, and what would cause us to be greater than 150, and look at how our scenarios fall in terms of whatever risk goal we define.

And the goal we can define maybe based on what people are doing already implicitly, but I have to figure that out. I'm not sure yet.

So, and then once we find the scenarios that might be important, we assess how they're doing in terms of the importance dimensions of looking at defense-in-depth, public confidence, incompleteness, and so on, and assess the need for supporting information and the value of imperfect information distribution that would be likely from whatever studies that are planned.

And I think one of the key things is how we could best use expert judgment elicitation, because

I think there is -- well, we talked a little bit about that. I'd like to see how we could decompose disagreements that are based on information versus interpretation of information.

And I'd like to see how the current DOE studies that are planned for the waste package and the drip shield and the igneous activity -- how they might do in terms of looking at the studies from this perspective, just because I think there has been a history of a lot of expensive studies that I don't know achieved what they originally wanted to achieve.

So anyway, well, that's where I'm going with this. I think I forgot like half of what I wanted to say, because it's very early. So I hope you can ask me questions that will remind me what I was supposed to say.

But I -- you know, thank you to everybody who has helped me over the summer and afterwards. I really, really appreciate it.

CHAIRMAN GARRICK: Okay. Thank you. We'll ask a few questions I think.

Ruth, do you have some?

MEMBER WEINER: I look like I have some. I have a couple of comments, actually. The first is you seem surprised that it was difficult to find scenarios and realizations, that the probability was very low that you would break the limit so to speak and get a very high dose.

Well, Yucca Mountain was selected because of some obvious external characteristics that it had that people thought it might make a good site. And I'm sure you can find sites that were not -- were eliminated from consideration where the probability would be very high.

This has to do with how these sites were selected in the first place, and I would suggest for your dissertation that you take a look at some of the history of the site selection. You probably have, but I think you probably -- you might look at it -- at the early environmental assessments, the five to three to one decision, and so on. That was one point.

The other point that I wanted to make -- and, again, I'm sure this will be in your dissertation -- is there were a lot of assumptions about solubility that you sort of sailed over and said, "Well, I'm just using neptunium-237." And I think that -- I hope that in your dissertation some of the uncertainties in that assumption are elucidated.

MS. GHOSH: Well, you know, the reason I had picked the neptunium-237 is to focus the studies.

We had talked about this at the beginning of the summer, and it was basically based on the NMSS studies that showed that it seems almost impossible to get a dose with anything other than neptunium-237, at least in the regulatory compliance period, which is why I sort of focused on that one.

And you're absolutely right -- in terms of finding the scenarios that could fail your criteria, yes, I knew going in it would be very difficult. Right? Extremely difficult. But I thought that if we can force the attributes into very conservative ranges that we might get some high probabilities of exceeding it.

So it's not that I'm surprised that there's no way to do some deterministic thing. I was surprised by the low probabilities for even the very high -- highly conservative assumptions.

MEMBER HORNBERGER: To follow up on that point, it strikes me that you have -- you face a difficult choice, because to get any kind of results into scenario classes you are almost certainly -- for example, your class of greater than 150 millirems, you're going to have to drive the model with extreme ranges only of the parameters. And what you lose there is, of course, you no longer have your risk basis, because you've walled it off.

An alternative way that I had thought that you might consider doing it is to say, well, okay, in the spirit of uncertainty and being bounded away from the limit, the dose limit specified in Part 63, you could ask the question -- just lower the calculated dose by a couple orders of magnitude, and you can say, "Okay. What drives it to be greater than .15 millirem? What drives it to be lower than 1.5 and above 1.5 millirems?"

And you might get a better feel for what a more realistic model would be doing in producing those doses.

MS. GHOSH: Yes. You're absolutely right, and I'm going to have to do that eventually. Because as I mentioned, I'm using the older, very conservative database. I think with the new one you can't even get a few realizations that exceed the limit. The only reason I had done that is because I always wanted to keep the decision context in mind.

So because that was the threshold -- but you're right that, yes, I think eventually I'll have to choose a much lower threshold. But then there's a real question of how we define the risk curve, because if we are so far from our decision threshold, why are we worrying about it? I guess there's still incompleteness.

But I guess my original idea is even though I knew I had to force the scenarios into very conservative assumptions, I wanted to find those things, so we could be confident that they were actually very conservative assumptions, because there's a lot that's left out of the analysis. And there just isn't resources to study everything.

So even though I knew that they were way out on the tail, I just wanted to first find the thing, so that we could convince ourselves that they really are far out on the tail. But, yes, I'm going to have to change the thresholds for binning.

CHAIRMAN GARRICK: Mike.

VICE CHAIRMAN RYAN: George, that's one of the observations I was going to make. And the second part is if you --

MEMBER HORNBERGER: I beat you to it.

VICE CHAIRMAN RYAN: You did.

(Laughter.)

And well, I might add.

If you're going to have this forced situation where you get realizations that exceed the dose, that will have a tendency to miscommunicate to the public. So you've -- I would tend to think about it this way, and I would suggest it might be helpful to you, is to try not to put all of the eggs into one basket.

In other words, if you're evaluating and exploring models, and components of models, and what's influencing what, at whatever dose level happens to make sense -- .15 or .015 millirem per year -- let that be its own assessment.

And then, if you're really trying to figure out what could exceed under a particular set of scenarios, or not exceed a given dose limit, that's also a related question, but really independent in a sense that you're focusing on dose rather than on processes. There's no real reason to couple those.

MS. GHOSH: Okay. You know --

VICE CHAIRMAN RYAN: At least in how you discuss them, because if you discuss them all at once people will assume if this process happens there's so many realizations above the dose curve, and that could be as ineffective of communicating and developing confidence as doing just a deterministic kind of job.

MS. GHOSH: Right. The thing is -- yes, I realize that's a very difficult issue, because on -- and, in fact, when the NRC had the model -- the uncertainty workshop over the summer, one of the DOE people said exactly that -- is that we're reluctant to show our bounding analyses, and so on, because people might just pick that up and say, "Well, look, it's possible, because you've proposed it."

So, I mean, I realize that's always an issue, and I want to repeat again that I don't -- whatever I'm proposing here, it's sort of supplementary analyses to a lot of the sensitivity analyses, and so on, that's already done.

But the thing is if we want to know like what might actually cause risk from the repository, there's no one place right now to get the answer. Right? I mean, we just -- so I realize there's a communication problem for communicating what is the meaning of such a study where we're looking at very extreme cases.

But I would like to have in one place an idea of what might actually cause the risk from the society, and then convince ourselves and everybody else, look how unlikely this is to happen. This is the only way we can figure out to even come close to exceeding our decision threshold. And it's just ridiculous to think that this is possible.

VICE CHAIRMAN RYAN: And if you keep all

of that in context, you can get there. It's just, you know, when you put together, it takes a little bit more art to communicate it completely, and I applaud your effort.

And that was my other comment. You've really jammed an awful lot of very good work in a very short period of time, and I congratulate you on your success with it.

MS. GHOSH: Yes. I'm really sorry. I wish I had more done. But by May --

(Laughter.)

CHAIRMAN GARRICK: Thanks.

I wanted to come back to the issue of uncertainty, and there's been a lot of thought given to how you characterize uncertainty in these kinds of analyses. And I want to ask you a couple of questions about whether or not some of these things have been considered.

When we first started doing very large risk assessments of nuclear powerplants, we were searching for a format that would be effective in communicating the risk. And one of the concepts that we came up with that was -- turned out to be quite powerful was kind of the probability of frequency concept.

And by that what we mean is that when you're dealing with the subject of likelihood, people talk about the likelihood of events, and so forth. The important issue here is: how do you represent likelihood? And there's maybe three ways to do that. One is to represent likelihood as the frequently of occurrence of a particular event. Another is the probability of the event.

And the third is a combination of the two -- is to admit that if you're looking for an event and its frequency, you know that there's uncertainty about that frequency. And the way you convey it is to embed that frequency in a probability distribution.

Well, that is a very -- that has been a very powerful format to follow. Now the question is: how would you do that in a repository situation?

Well, you'd probably do it the same way that it's been done in other applications. You think of it in the context of a thought experiment. You ask yourself if this went on for millions and millions of years, and you looked at a particular time interval, what's the frequency of occurrence of a certain damage level? And damage level can be anything from fatalities and injuries to dose.

And so, but this really then allowed us to characterize the risk in the context of a complementary cumulative distribution function format with frequency being the ordinate and damage being the abscissa and probability being the variable.

And so what we also found as far as the whole business of how you present this to not confuse the public is that you tend to collapse the percentiles into a few critically important ones, or at least the kinds of percentiles that you see in the risk community, such as the five percentile, the 50 percent, and the 95 percent, and maybe you want to put the mean in there, too.

So if you do that, then clearly you don't get into these dilemmas of having these horsetails that often go above a limit, because even in the one you've shown the 95 percentile would be well below the limit. It would be below the limit.

So that's one thought is that you -- when you present the material, you present it in the context of specific percentiles that are pretty characteristic of what we're used to, such as the 90 percent interval.

But I like the idea of presenting the risk in that form, in the form as you did at the outset here. Namely, you just put it down the way it is with the kind of adjustments I've just alluded to as possibilities.

I like that better than, in fact, the way it's being done, which is the peak of the means. It's just to present the CCDF of the dose of Yucca Mountain based on perhaps a five percentile, 50 percentile, and a 95 percentile, and then you -- it's pretty obvious when you impose the limit line on that where you stand.

So I'm curious as to whether or not you are thinking in your work about the fundamental issue of uncertainty and how to characterize it. And I made that speech because experience has indicated that this is -- there is at least one very powerful way to represent it -- through the probability of frequency idea. And I just wondered if you had considered that.

MS. GHOSH: Yes, yes. Thank you for bringing that up. I've definitely thought about it a lot. And one of the things that bothers me a little bit is that either if you look at a reactor PRA, which typically you look at some measure like the core damage frequency, or the width PA, and you look at the curves that were generated in the summary measures, there is a separation of aleatory and epistemic uncertainty. Right?

Because with the -- in a tree, through a reactor, you are looking at events that might happen that one can think of as aleatory uncertainty. And the epistemic uncertainty comes from the -- you don't know about the failure rates for the various things that have to occur.

And the same for WIPP. When we had the family of curves for WIPP, each curve represented epistemic uncertainty, and the actual CCDF was for aleatory uncertainty for their dominant human intrusion scenario.

Now, for Yucca Mountain we lumped the aleatory and the epistemic uncertainty into one dose history, right? Because our parameters are actually capturing all kinds of things, not just uncertainty and assumptions. So there are also chance events that are lumped in, so everything is lumped into the dose history.

So there is -- overall there is less information, right, about the type of uncertainty that's represented in the curve. And I guess there's not a lot to do about that, but, of course, I like your idea about when we represent the CCDF to look at some percentiles rather than just having one.

I think it's a good idea, and I'll

definitely do that for my -- as I progress. But that's definitely one thing that has bothered me about the assessment for Yucca Mountain. I guess maybe nobody else cares, nobody else is concerned about it. But we've lumped more into each, you know, realization than, for example, for WIPP, because with the WIPP the criterion was much simpler. So you are able to represent the different types of uncertainty.

CHAIRMAN GARRICK: Well, I think people characterize it a little different. Rather than talking about epistemic and aleatory, they talk about information uncertainty and modeling uncertainty. But it's the same thing. And I think that those two issues probably have to be treated somewhat separately.

You know, the business of information uncertainty is very much more advanced in terms of a science than is the business of modeling uncertainty. And the way that often it's been done with respect to the modeling uncertainty is to apply different physical models to the same problem and see what kind of variations or perturbations you get. But there are ways of doing -- at least getting some handle on both of those kinds of uncertainty.

Any other questions? Yes.

MEMBER HORNBERGER: I have just a couple nitty-gritty things that I want to explore, then, okay? So if I look at your slide on crude estimates for scenario probabilities, okay, it seems to me that there's a lot, shall I say, hidden in here. It's not entirely obvious. So the first thing is, if I understand this correctly, you had to make this decision that one, two, three, four, five, six factors or parameters in the model were all that was important.

MS. GHOSH: Yes. But as I mentioned, it's not true --

MEMBER HORNBERGER: No, I know. But you had to make that assumption, right?

MS. GHOSH: Right. That is my initial assumption, yes.

MEMBER HORNBERGER: That is your initial assumption. Okay. So now my question, then, when I look at that table is, I don't know why you didn't fill in numbers for each one of those factors to calculate the exceedence probability.

You know, even if infiltration in the 101 dose was at the one percent level --

MS. GHOSH: Right.

MEMBER HORNBERGER: -- because it strikes me that sort of your expected exceedence probability for any realization is one over 2^6 in this case, if you have six parameters. Is it not?

MS. GHOSH: Say that last thing again.

MEMBER HORNBERGER: Okay. So the expected value of your exceedence probability --

MS. GHOSH: Yes.

MEMBER HORNBERGER: -- if you have a model for six parameters that are all important is one over 2^6. Half of your realizations are going to be above that value, right?

MS. GHOSH: Yes. Oh, right.

MEMBER HORNBERGER: So you have an expected value of this last column, and so your comparison -- I would think that you would want to fill in that whole table. Small point.

But it's also clear, though, that if you did this and anticipated that all 100 parameters were important, then your expected exceedence probability would be one over 2^{100}, which is now a really small number. And so that's why it's hard to pull apart this for a very complicated model.

CHAIRMAN GARRICK: Yes. I think that's one of the things I was alluding to, I wasn't getting to, is I think it's maybe very dangerous to address each realization in this context. I think that this is a case where lumping them would make the problem have a lot more physical meaning, because it's a kind of -- the realization -- when you get down to five or six realizations, it's almost an artifact of the calculation rather than reality. And so --

MS. GHOSH: Can I propose something?

CHAIRMAN GARRICK: Yes.

MS. GHOSH: And you tell me what you think. So I -- first, I have to increase the number of realizations. I was thinking maybe -- let's say we start with 1,000 or something, right? And I have to do some multiple sets of 1,000 realizations, if I want to probe the sample spaces for various attributes, right?

So let's say we pick things that make sense to look at together, like the initial waste package rate, failure rate, with the infiltration and how much water is actually getting there, because that gives us our source term coming out of the waste package.

So if we force those, let's say, into some -- in the 25th percentile range, let's say, for those things, and let everything else vary as it is in the analysis, and I should get something about -- first, I will get a dose distribution with some exceedence probabilities.

And as I said before, you're still going to have a lot of cases where you don't exceed the dose, but what you would end up with is maybe a couple of end states, where you're going to exceed like 100 millirem or 15 millirem when you're in the sample space. Right?

So now, would it be improper to maybe define this class of scenarios and say that if we have this class of scenarios we -- basically, to build a CCDF for the class of scenarios. So, you know, we found that in 40 percent of the realizations we still have a zero dose, but we may exceed 100 with this probability based on this sample space.

Yes. Because, I mean, that's what I was originally trying to do with the nine runs, and it's just not enough to do that. Yes. Because, I mean, the point of doing this, and I take your point, but I wanted to see for this particular history of how the repository might have evolved, how can we explain what we actually found at the end? And I didn't include the parameters where it was below the 50th percentile, because I made the assumption that it doesn't explain why we got the dose, because it could have been any value.

MEMBER HORNBERGER: It wouldn't change your numbers very much. Your right-hand column would not be changed very much, if you think about it. Because what's affecting your right-hand column are the 99 percent and the 98 percent. Anything less than that isn't going to affect it very much, as you saw for the one that came out to be 1.77 times 10^{-2}.

CHAIRMAN GARRICK: Yes.

MEMBER HORNBERGER: Those things don't make the number low. You're basically talking about factors of two rather than 10.

Okay. Could I go on to something else?

CHAIRMAN GARRICK: Yes, sure.

MEMBER HORNBERGER: So the other thing that intrigued me is your hypothetical risk goals. And I understand the WIPP CCDF and the nice way that the EPA had that standard. But I'm a little confused that -- well, to me, I'm not sold that these hypothetical goals that you put forward make sense. Okay? And let me tell you my reasoning. Okay?

This is essentially a cumulative distribution, and so the slopes of the cumulative distribution give us density. Okay? So your top goal indicates that -- to me, that you have a zero probability density of getting any dose between 15 and 150 millirems. Okay?

Now, the only way that makes sense to me is if you extend that line horizontally out to infinity, which means that your hypothetical risk goal is to have a zero probability that you will have any dose exceeding the limit, which --

MS. GHOSH: Wait. Why do you see there's a zero probability? There is a 10 percent chance for the --

MEMBER HORNBERGER: There's a 10 percent chance for the limit.

MS. GHOSH: Yes.

MEMBER HORNBERGER: Okay, at 15 millirems.

MS. GHOSH: Yes.

MEMBER HORNBERGER: And it's a 10 percent chance at 150. So how many additional ones have you accumulated? Zero.

If you take a cumulative distribution function, the slope of that cumulative distribution function is the probability density. Okay? The slope of that cumulative distribution that you have there is zero.

MS. GHOSH: I'm sorry. I just took this straight from the WIPP, and the line that I draw is irrelevant. I think the main points are they didn't want -- that they didn't want to overprescribe, you know, what the goals are. There is just two points, mainly the one at the bottom and the --

MEMBER HORNBERGER: I know. And I know WIPP is a release fraction, so it's a different thing.

MS. GHOSH: Yes.

MEMBER HORNBERGER: And that's my point. My point is that what I think you should do is think about whether that translates very nicely over to a dose standard or not, because your other one the slope I don't like any better either as a risk goal, because what it indicates is that you have an equal chance of getting any dose between 15 and 150, which doesn't seem to make a lot of sense to me either.

MS. GHOSH: Okay. Yes. I will think about that.

CHAIRMAN GARRICK: I think a much more straightforward approach is just to prevent the CCDF on the basis of the scenarios. Now, at the end of each scenario you have a PDF, a probability density function, and each scenario has a different end state.

So you take those scenarios and you organize them in the order of increasing damage, and then you accumulate them from the bottom up. And that gives you your cumulative distribution. And just present that distribution as it is, but have -- with a clear trail from the scenarios.

MEMBER WEINER: Can I make one additional point?

CHAIRMAN GARRICK: Yes.

MEMBER WEINER: If you read 40 CFR Part 194, you will find that there is also a dose standard for the WIPP, although the one that was applied was this. And a more appropriate comparison for -- to Yucca Mountain would be to use that dose standard, because it -- as George has pointed out, this -- the probability of one and probability of .001, that was specifically for the release standard.

So I would encourage you to go back to 40 CFR 194 and redo that. Look at the -- well, there are -- I've forgotten what the section number is. But 40 CFR Part 194 was the regulation that EPA wrote for the WIPP. That was the one that we had to write the compliance certification application to. And that also provides a dose standard. I think that might be a more logical one to use.

MS. GHOSH: Okay, yes. Thank you. I forgot about the dose for WIPP. So I'll look it up.

CHAIRMAN GARRICK: Excellent. Go ahead,

George.

MEMBER HORNBERGER: This is really -- it's not even a nitpick. It's just -- I put a question mark here, and Neil can probably answer the question for us. Do you know how many volcanologists have been killed in eruptions?

MS. GHOSH: Oh, is it high? Because I didn't know. Is it true? Yes? Oh, everybody is nodding. Okay. So I'm --

MR. CAMPBELL: It's one of the riskiest jobs in all of earth sciences.

MS. GHOSH: Is it worse than astronauts? Yes? Okay. Well, thank you for pointing that out, because obviously I'm naive. I had no idea. I was young when Mount St. Helens happened, so this was really impressed in my brain, that this poor scientist who loved his work died, you know, in the thing he was studying.

CHAIRMAN GARRICK: I would be -- I wouldn't want to have the record be too categorically sure about that it's worse than astronauts, because the astronaut sample is very small, and the incidents are quite high, especially if you consider all of the programs.

But anyway, we are very impressed with

what you're doing and --

MS. GHOSH: Can I ask you one last question?

CHAIRMAN GARRICK: Yes. We have -- excuse me. We have some other questions.

MR. McCARTIN: I just want to -- just to make an observation, but also your Table 3 -- I'd like to compliment you on it. I think it's a very interesting way to present some information to make people think. And I'll just point to one thing that I have to go back and scratch my head on.

I have looked at the source code for LHS sampling, and there is a lot of code in there to not introduce correlation -- unspecified correlation. And when I look at your first vector there, the largest one, there's three parameters that are at their 97th or higher percentile. Those three are uncorrelated, and it's in a sample of 200. You could get three parameters at that high of a value. It's kind of fascinating.

Now, it is random sampling, so it doesn't mean it can't happen. However, it -- I'm going to go back and just try some statistical experiments with our input --

MEMBER HORNBERGER: It would be really

amazing if it were random sampling. That's why you get --

MR. McCARTIN: Yes. But even -- it does not try to do this. You can get it, but I'll tell you it would be interesting in 200 vectors to get three uncorrelated parameters to be almost at their highest value. I'd like to do, as you suggested you might do, a 1,000 vector sample.

MEMBER HORNBERGER: How many are you sampling?

MR. McCARTIN: Well, and that's why I'm wondering if LHS gets confused with, you know, the approximately 2- to 300 sample parameters.

MEMBER HORNBERGER: You are sampling that many?

MR. McCARTIN: Yes. Yes. It will just be interesting to do something -- some experiments, some statistical experiments. I mean, you might do 1,000 vectors and not get more than one that lines up that well. I mean, in any given set you can get some rare events. It doesn't mean that every 200 vectors sample you'll get that, but it is an interesting result.

I'm not sure what it means, but, once again, it's the benefit of here's a different way of looking at things. I think there is, as Dr.

Hornberger indicated, there's a lot of information here. But that's one of the -- I compliment you on your -- the work that you spend here. That's one of the things we certainly hope to get when we have -- intern work is stuff that makes us think.

MS. GHOSH: Well, I based it on your study, so thank you.

(Laughter.)

MR. McCARTIN: Well, I didn't --

MS. GHOSH: I was just trying to quantify what you had already found, but, you know -- am I allowed to ask you guys a question? Because I just want to -- okay, because I really want to think about how to have a risk goal in terms of the CCDF. And so, Dr. Hornberger, it seems like you don't -- is it that you don't like the idea at all, or you don't like the specific examples that I showed?

MEMBER HORNBERGER: The latter.

MS. GHOSH: The latter.

MEMBER HORNBERGER: I guess what I'm suggesting is you need to think about it as to whether your hypothetical goals on a CCDF for a dose standard make sense. Okay. And I'm not judging -- I'm not saying that you might not be able to come up with a hypothetical goal, but I don't think it's quite as simple as just transferring the WIPP release fraction goal over onto a dose standard.

MS. GHOSH: Right, okay. I'll definitely think about it. Thanks.

CHAIRMAN GARRICK: I think, again, it's more important to be able to present the dose as to what it is. You know, and if it's -- and if you have a dose standard, as we do in Yucca Mountain, then you know exactly where you stand. So I --

VICE CHAIRMAN RYAN: That's kind of my point, John. Is there really a need to have a surrogate for dose? You know, in terms of a risk coefficient or some other risk goal? Why don't we just use the dose?

MS. GHOSH: Oh, no, no, no, no. It's just that if we look at just the mean, we don't look at the whole spectrum of risk. That was my point in trying to build some way to look at the whole distribution.

VICE CHAIRMAN RYAN: Yes.

MS. GHOSH: That's why I brought it up. Of course, dose is what --

VICE CHAIRMAN RYAN: Sure.

MS. GHOSH: -- is causing the risk. Yes.

VICE CHAIRMAN RYAN: But, I mean, if you follow Dr. Garrick's comment and just, you know, put all of the realizations up against that dose -- and even if it's in some sort of a forced way, so you can see what's happening, you know, from a model perspective, that's useful. But presenting the dose is always meaningful to me.

CHAIRMAN GARRICK: Yes. Phil, go ahead.

MR. JUSTUS: A perspective on your Mount St. Helens example, and I am not doubting that it was a surprise certainly to the 60 people who died from the horizontal blast effect.

But the alternative perspective is that from a regulatory point of view Mount St. Helens was a -- was the agency's first success stories in designing a nuclear powerplant to mitigate the effects of a volcanic eruption. I'm speaking of the Trojan plant.

And the U.S. Geological Survey did correctly determine the weak side of a future Mount St. Helens blast. They correctly determined the vertical plume and the direction of dispersion of it. They correctly predicted that the greatest hazard to the Trojan plant would be from the mud flows coming down the flanks of the volcano into the Columbia River, clogging their intake system with silt, mud, and sandy particles, and such.

And there would be some ash, and the HEPA filter system and other filters would need to be employed to withstand the small amount of ash from a future eruption. All of that happened within the design basis for that volcanic event.

Just another perspective on -- it was a surprise, but yet it wasn't from another point of view.

CHAIRMAN GARRICK: Any other comments, questions, discussion? Well, thank you very much. Good luck in your activities from now until May.

MS. GHOSH: Thank you. Thank you all very much. I really appreciate your helpful comments.

CHAIRMAN GARRICK: Any other comments from the committee or staff before we adjourn for lunch? All right. I think we'll adjourn for lunch, and we'll resume at 1:00.

(Whereupon, at 11:54 a.m., the proceedings in the foregoing matter went off the record for a lunch break.)

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 Nuclear Waste 146th Meeting

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.

Jennifer Rosario
Official Reporter
Neal R. Gross & Co., Inc.

Risk-Informed Uncertainty Studies for the Yucca Mountain HLW Repository Program

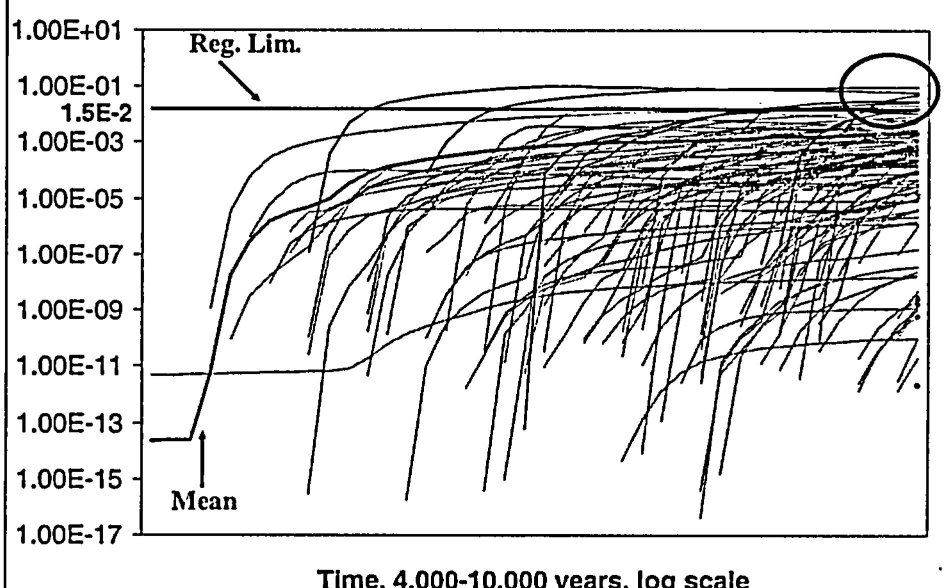
10/21/03

Tina Ghosh

ACNW Summer 2003 Intern

PhD Candidate, MIT Nuclear Engineering Department

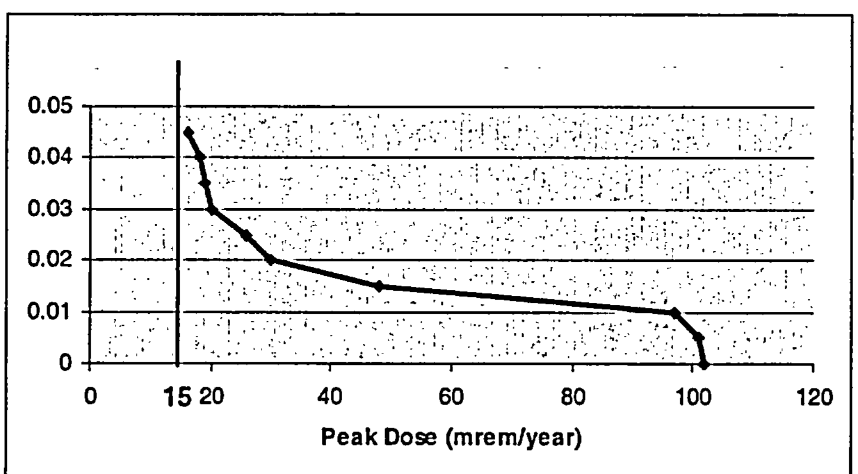
Np-237 Dose (rem/yr) for 200-Realization Base Case



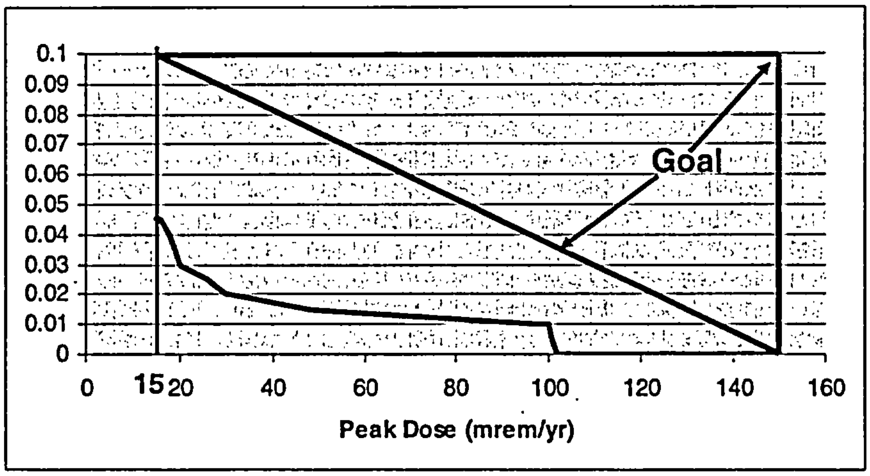
Crude Estimates for Scenario Probabilities Based on TPA 4.1 Database

Dose (mrem/yr)	Infil-tration	WPFlow Factor	WP_ defect	Solub-ility	Kd CHv	Rf-Alluv.	~Exceed Prob.
101		99%	98%	97%	64%		2.14E-06
97		97%	80%	96%	64%		8.64E-05
48	97%	95%	69%		55%	92%	1.67E-05
30		94%	96%	89%		60%	1.06E-04
26	82%	88%	93%	80%			3.02E-04
20	60%	87%				66%	1.77E-02
19	90%	74%		94%	95%		7.18E-05
18	91%	71%	81%		81%	89%	5.45E-04
16			63%		99.8%		7.41E-04

CCDF for 9 Realizations of Interest

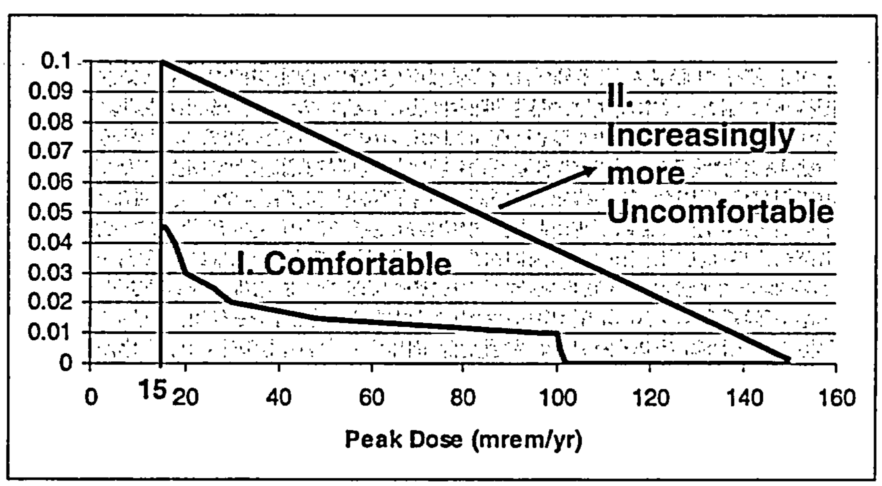


CCDF in Relation to Hypothetical Risk Goal



NRC Advisory Committee on Nuclear Waste 5 Massachusetts Institute of Technology

Risk Goals and Comfort Zones



NRC Advisory Committee on Nuclear Waste 6 Massachusetts Institute of Technology

Model Uncertainty Can Complicate Probability Estimates

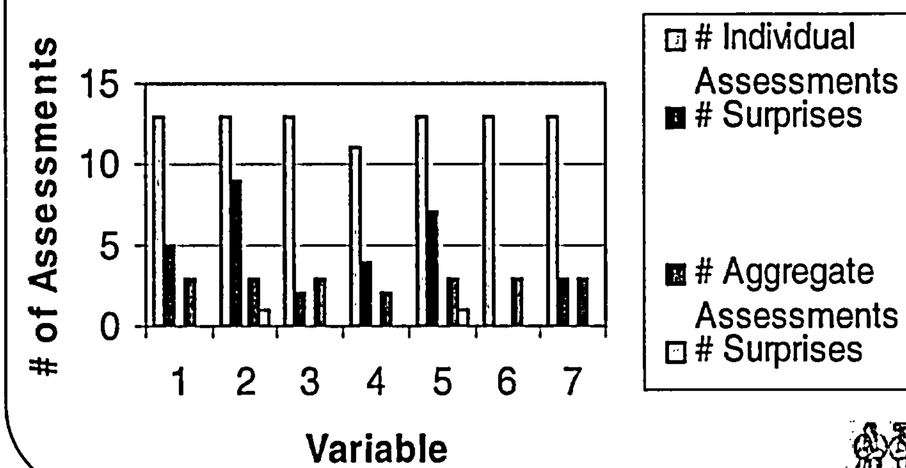
- Conceptual model uncertainty
  - ➢ Alternate models available or one model thought to be weak
  - ➢ Dependencies among variables and/or coupling between processes not considered or omitted
  - ➢ Inconsistencies
- Uncertainty introduced through model abstraction process
- Biases, overconfidence
- We don't know what we don't know – incompleteness

Incomplete-ness

- RIP David Johnston, USGS vulcanologist, MT. St. Helens, 5/18/80
- The magnitude (F=~24MT bomb), direction (sideways instead of up), and consequences of the eruption stunned the entire scientific community

"Surprises" in Expert Assessments of Seven Nuclear Accident Variables

(33% surprise for indiv., 10% for aggregate)



NRC Advisory Committee on Nuclear Waste 9 Massachusetts Institute of Technology

Importance Ranking Dimensions

- Assessed risk
- Confidence in these risk estimates and likely nature of any biases
- Incompleteness
- Defense-in-depth
- Public concerns and confidence

NRC Advisory Committee on Nuclear Waste 10 Massachusetts Institute of Technology

Confidence Building -- Conservative Biases

- TPA 4.1 has many conservative assumptions about Engineered Barrier System (EBS), examples:
  - WP failure is binary and does not take credit for cladding; DOE assumes ~90% protection, although NEA/IAEA review noted that basket corrosion (left out of analysis) may affect cladding
  - Assumes bathtub immersion model for radionuclide release from waste form, even if there is not enough water for such an effect
- Another example: average dilution volume calculated by TPA was 5% of the EPA's representative volume
- Potential reduction by 2 orders of magnitude

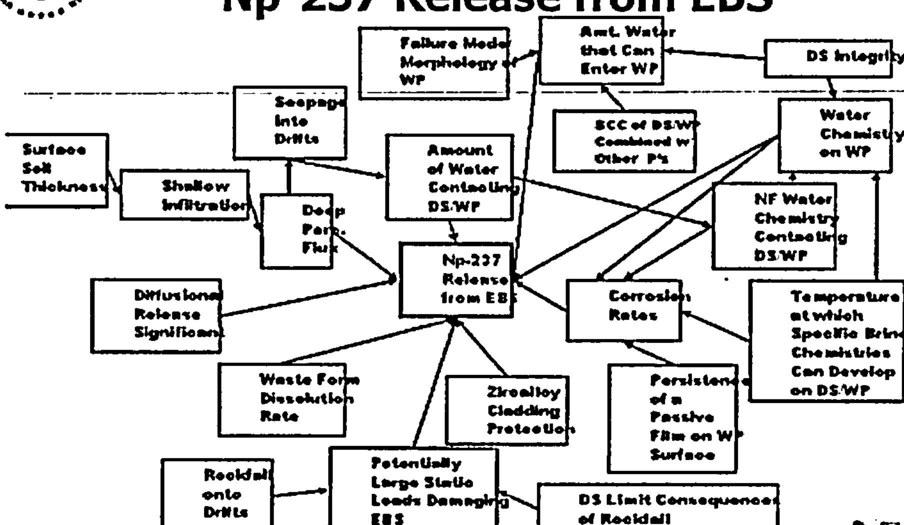
Potential Adverse Effects of Incompleteness

- Assumed no failures due to general corrosion before 40,000 years
- Possible groundwater fast(er) paths?
  - Look at past cases to get an idea of order of magnitudes possible
    - e.g., LLW disposal sites, contamination travel at most DOE complex sites, possibly Chernobyl land contamination transport; natural analogs
- For example, if 0.1% of nuclides can travel 10x faster, how much larger could the dose be?
  - Take weighted fraction of UZ release (which on average is 3X higher than SZ release) for a crude estimate; possible increase ~3%; most likely not significant

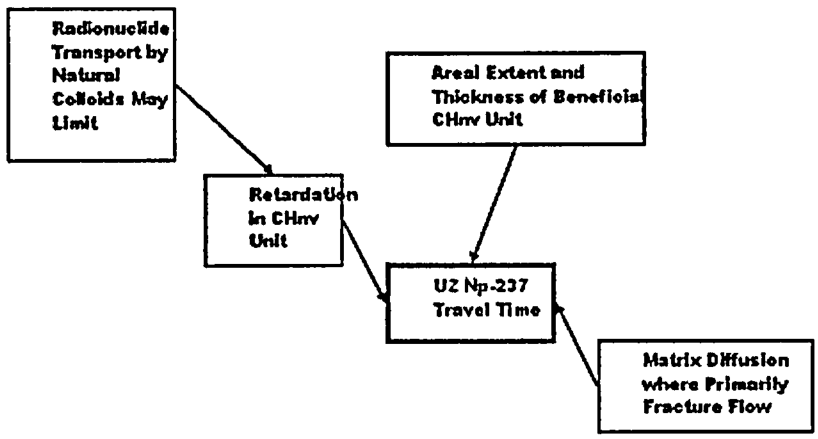
Incompleteness Estimate Based on Disagreement

- We can consider the DOE and NRC as two expert communities
- DOE and NRC disagree about the relative capability of the UZ vs. SZ to attenuate radionuclide plume -- why?
  - ➢ Conservative assumptions made for simplicity and to help focus available resources?
  - ➢ Legitimate disagreement because of different beliefs about dominant processes?

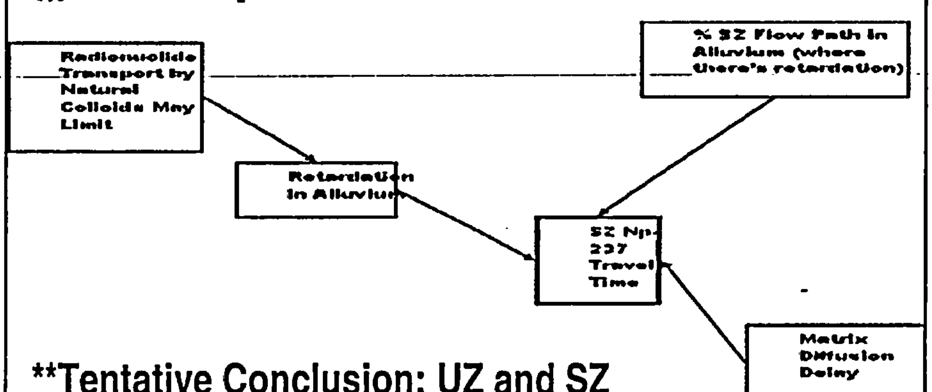
Influence Diagram of KTIs Affecting Np-237 Release from EBS



Influence Diagram of KTIs Affecting Np-237 Release from the UZ



Influence Diagram of KTIs Affecting Np-237 Release from the SZ



**Tentative Conclusion: UZ and SZ capability should be independent of EBS capability

Confidence Building -- Dependencies and Coupled Effects

- Infiltration/water flow into WP and waste form release rates from EBS may have coupled and synergistic effects
  - ➢ Potential localized corrosion due to salt deposits; related to water cycling in HTOM
  - ➢ Evaluate in terms of requirements to prevent "failure", rather than e.g., sensitivity to 95th percentile
- These coupled effects may lead to a different assessment of individual EBS barrier capabilities (than assessment based on radionuclide concentration attenuation after neutralization of prior barriers)

Defense-In-Depth Measures

- How far are we into the "comfortable" region? (a safety margin?)
- Is there any single assumption that, if wrong, could defeat the system? (similar to single-failure criterion for reactors)
- How confident are we in the probability assessed for the scenarios with the fewest elements/processes? (measure of diversity goal)

Confidence Building -- Qualitative Aspects

- What are the technical bases for the important assumptions?
  - ➢Are these based on empirical evidence?
  - ➢Is the evidence (e.g., laboratory experiments) directly applicable, or is extrapolation and inference involved?
    - WP corrosion studies a good candidate for investigation
- What kind of peer review was used?
- Is the environment conducive to analysts raising concerns? (culture of decision-making organizations)

Public Debate of Technical Issues

- Water in Yucca Mountain -- where does it come from, where does it go?
- Igneous activity -- will the volcanoes remain dormant? If not, what could be the consequences?
- Criticality -- potential success story: how was this issue resolved?
- What are the corrosion rates and mechanisms for Waste Package and Drip Shield?

Assessing the Need for, and Value of, Further Studies

- Should take into account the value of imperfect information (VII) that would be obtained through the studies
  - ➢The expected VII is probably not good enough; may be worthwhile to develop a distribution
- Should consider all of the importance dimensions
- Current DOE efforts may be going in this direction, but focus and methods are slightly different

Future Work

- Define end states (e.g., 15-150 mrem Np-237 dose, >150 mrem)
  - ➢ Risk goals based on current practice?
- Find classes of scenarios that lead to such end states (probability>0.5?)
- Assess importance dimensions for these scenarios
- Assess need for supporting information, and the VII distribution likely from planned studies
  - ➢How best to use expert judgment elicitation?
  - ➢Evaluate DOE studies planned for WP/DS corrosion, and igneous activity?

Acknowledgements

- ACNW members and ACNW/ACRS staff
- Tim McCartin, David Esh, and Richard Codell (NRC's NMSS)
- George Apostolakis (MIT), thesis supervisor

Back-up Slides

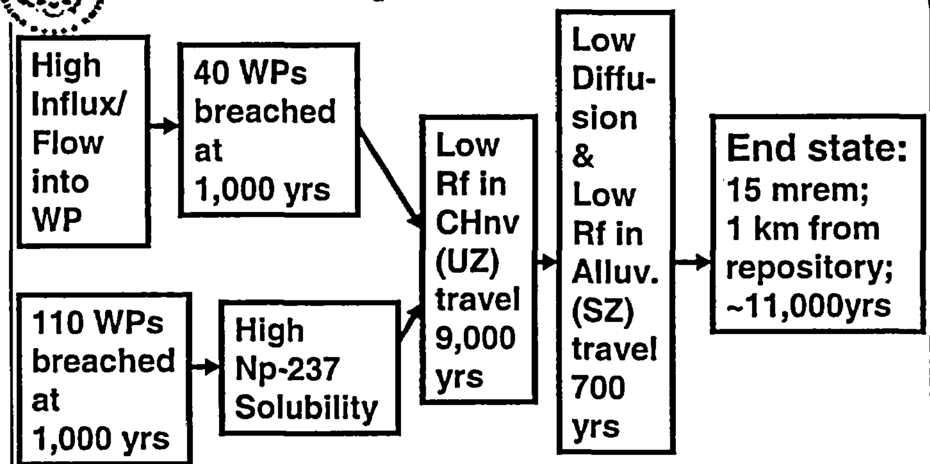
Key EPA Guidance in 40 CFR 197

- Accessible Environment starts at (on south) 36º 40' 13.6661" north latitude in predominant direction of GW flow, or 5 km in any other direction from repository footprint
- Representative Volume = 9.77 x 10^8 gal/yr
- Dose limit = 15 mrem/year to RMEI
- GW dose limit (from β and γ emitters) = 4 mrem/yr
- GW activity limit for most α emitters = 15 pCi/L

Parameter Values for Sample Runs Scrutinized

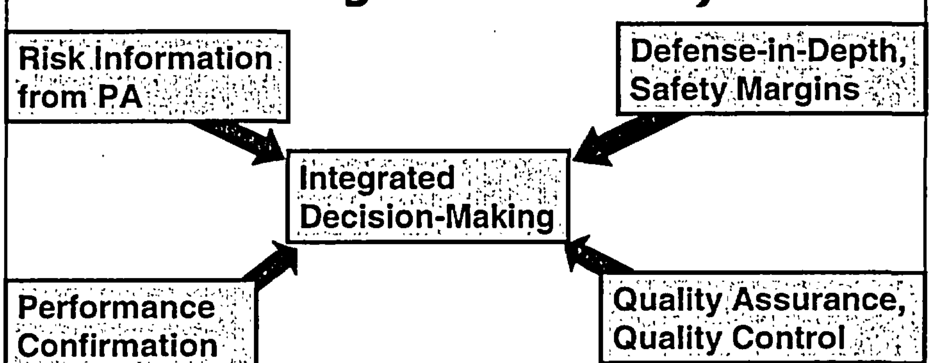
			Key Variables and their index values						
Sampled	Dose (mrem)		AAMAI@5 (1)	WPFlowMF (60)	WP-Def% (62)	Solb-Np (67)	MKDChAn (14)	ARDSAVNp (225)	WPRRG@10 (238)
64	100.6		5.75E+00	2.97E+02	9.77E-03	1.38E-01	2.60E-02	1.79E+02	8.83E+04
14	87.098		4.72E+00	1.49E+02	8.01E-03	1.28E-01	2.61E-02	4.80E+01	4.47E+04
134	47.593		1.27E+01	8.29E+01	6.83E-03	2.82E-02	2.84E-02	8.33E+00	2.53E+05
189	29.767		7.11E+00	8.33E+01	9.99E-03	8.23E-02	3.02E-02	4.51E+01	1.84E+04
186	26.471		1.14E+01	4.24E+01	9.32E-03	5.58E-02	3.89E-02	2.69E+02	1.40E+05
103	20.097		8.40E+00	2.82E+01	4.10E-03	3.01E-02	3.15E-02	3.57E+01	2.53E+05
32	18.86		1.21E+01	1.89E+01	5.17E-03	1.12E-01	1.83E-02	1.44E+02	2.30E+05
163	18.413		1.22E+01	1.48E+01	3.78E-03	3.35E-02	2.18E-02	1.17E+01	4.78E+04
199	16.097		8.65E+00	5.44E+00	6.31E-03	1.84E-02	1.08E-02	1.32E+02	1.07E+05
		sample mean	8.34E+00	8.35E+01	7.00E-03	6.83E-02	2.35E-02	8.85E+01	1.34E+05
		total range	4-13 mm/yr	3.15e-2->1.05e3	1.0E-4->1.0E-2	1.2e-3 3.4e-2, 2.4e-1 (kg m^3)	9.91E-3->8.84E-2 (m^3/kg)	1-3.9E3	1.5e4->2.84e5 (g/m-day)
		distribution	uniform	lognormal	uniform	logtriangular	lognormal	lognormal	uniform
		distribution	8.5	6.75	5.00E-03	1.71E-02	0.0296	62.5	1.49E+05
		median							

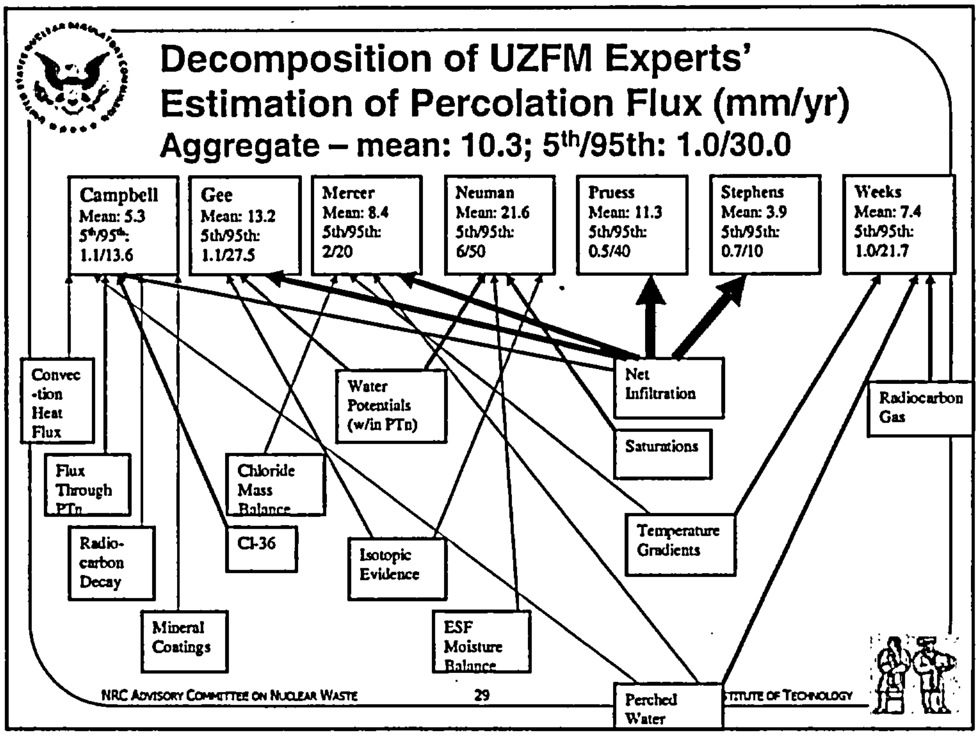
Example of Scenarios

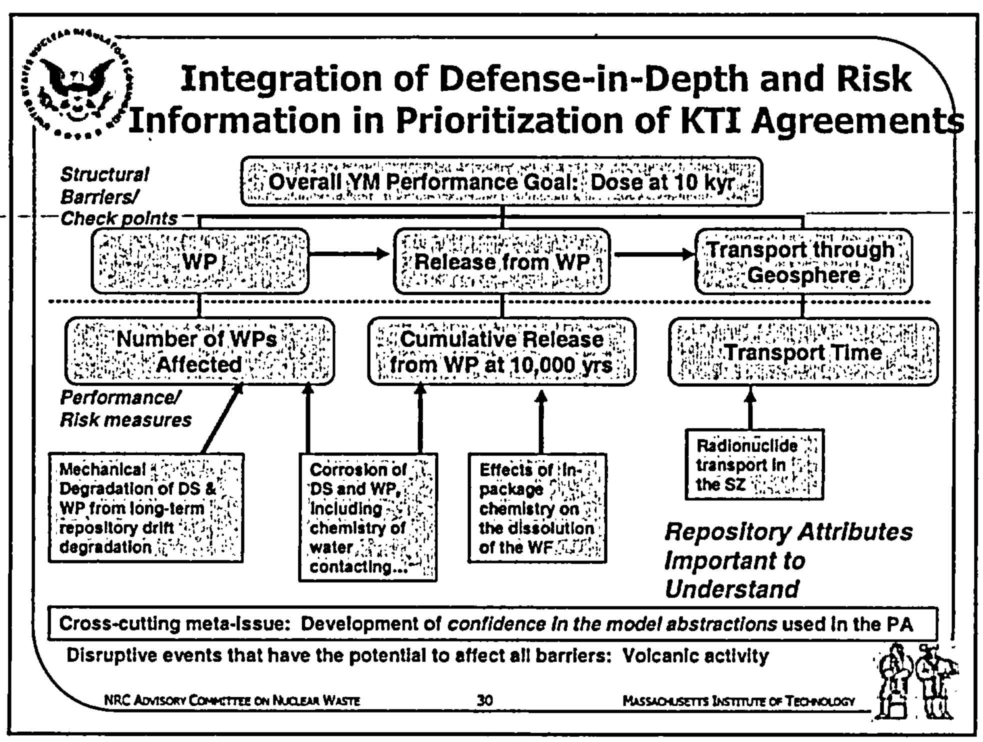


Based on NRC NMSS' Tim McCartin's 3/26/03 Presentation to the ACNW

Risk-Informed Integrated Decision-Making Framework (adapted for YM from US NRC Reg. Guide 1.174)




Decomposition of UZFM Experts' Estimation of Percolation Flux (mm/yr)
Aggregate – mean: 10.3; 5th/95th: 1.0/30.0
Campbell Mean: 5.3 5th/95th: 1.1/13.6
Gee Mean: 13.2 5th/95th: 1.1/27.5
Mercer Mean: 8.4 5th/95th: 2/20
Neuman Mean: 21.6 5th/95th: 6/50
Pruess Mean: 11.3 5th/95th: 0.5/40
Stephens Mean: 3.9 5th/95th: 0.7/10
Weeks Mean: 7.4 5th/95th: 1.0/21.7
Convec-tion Heat Flux
Flux Through PTn
Radio-carbon Decay
Mineral Coatings
Chloride Mass Balance
Cl-36
Water Potentials (w/in PTn)
Isotopic Evidence
ESF Moisture Balance
Net Infiltration
Saturations
Temperature Gradients
Radiocarbon Gas
Perched Water
NRC Advisory Committee on Nuclear Waste
29


Integration of Defense-in-Depth and Risk Information in Prioritization of KTI Agreements
Structural Barriers/ Check points
Overall YM Performance Goal: Dose at 10 kyr
WP
Release from WP
Transport through Geosphere
Number of WPs Affected
Cumulative Release from WP at 10,000 yrs
Transport Time
Performance/ Risk measures
Mechanical Degradation of DS & WP from long-term repository drift degradation
Corrosion of DS and WP, including chemistry of water contacting...
Effects of in-package chemistry on the dissolution of the WF
Radionuclide transport in the SZ
Repository Attributes Important to Understand
Cross-cutting meta-issue: Development of confidence in the model abstractions used in the PA
Disruptive events that have the potential to affect all barriers: Volcanic activity
NRC Advisory Committee on Nuclear Waste
30
Massachusetts Institute of Technology