UNITED STATES

NUCLEAR WASTE TECHNICAL REVIEW BOARD

FALL BOARD MEETING

DEVELOPING A REPOSITORY SAFETY STRATEGY
WITH SPECIAL ATTENTION TO MODEL VALIDATION

Radisson Plaza Old Town Hotel
Alexandria, Virginia

Wednesday, September 15, 1999

NWTRB BOARD MEMBERS PRESENT

Mr. John W. Arendt
Dr. Daniel B. Bullen
Dr. Jared L. Cohon, Chair, NWTRB
Dr. Paul P. Craig, Afternoon Session Chair
Dr. Debra S. Knopman
Dr. Priscilla P. Nelson
Dr. Richard R. Parizek, Morning Session Chair
Dr. Donald Runnells
Dr. Alberto A. Sagüés, Roundtable Discussion Chair
Dr. Jeffrey J. Wong

SENIOR PROFESSIONAL STAFF

Dr. Carl Di Bella
Dr. Daniel Fehringer
Dr. Daniel Metlay
Dr. Leon Reiter

CONSULTANTS

Roger Newman
Naomi Oreskes

NWTRB STAFF

Dr. William D. Barnard, Executive Director
Michael Carroll, Director of Administration
Karyn Severson, Congressional Liaison
Vicki Reich, Librarian
Ayako Kurihara, Editor
Paula Alford, External Affairs
Linda Hiatt, Management Analyst
Linda Coultry, Staff Assistant

I N D E X

P R O C E E D I N G S

9:00 a.m.

PARIZEK: We'll begin in about 30 seconds for the morning session. We're going to continue today on repository safety strategy. There will be two presenters this morning, the Testing and Analysis for Site Recommendation by Jean Younker, and then Introduction to Validation by Bob Andrews. Then we'll have a question and discussion period. There will be a break, and then a period for public comment.

This afternoon, there will be two presentations that relate to Model Validation, one being the unsaturated zone example, and the other the waste package example. Those will be interesting because I guess these would be guinea pigs of two programs that have gotten pretty far into this process on model development.

Then there will a roundtable discussion this afternoon starting at 3:00, and if you have the agenda for the program, you'll see a number of people are listed there. There's been some modification. Norm Christensen, who was going to be the Chair for the program, because of the hurricane has left, and he has to take care of some university business in the wake of that arising hurricane. So he will not be with us, but we have Alberto Sagüés, who will be the Chairman in his place, and then I've been asked also to be present for that discussion. And at the end, there will be again closing remarks and opportunity for some public comment.

Now, the first presenter of the morning would be Jean Younker. She's obviously well known to everybody attends these meetings on a regular basis. But while she's getting ready for her presentation and coming up, I just want to say that she did her Bachelor's Degree in Physical Science and a Master's Degree in Physical Science and Geology, and a Doctorate in Geology at Michigan State University, has had important activities with the program for a number of years. Prior to getting in this part of the program, she was at Lawrence Livermore National Lab, and held various academic position in her earlier part of this effort, and she has major responsibilities with the program at the present time.

So, Jean, we look forward to your remarks.

YOUNKER: Thank you. Let me say good morning to everyone, and say that this presentation is a follow-on to what you heard from Mark Peters yesterday, where Mark gave you an indication of what kind of results we had that are being used as pretty much direct input to the first revision, what we call Rev. 0 of our analysis and modern reports that support the preparation of the overall technical basis for site recommendation.

What this one does is picks up with that testing that continues on over the next 18 months, some of which will perhaps provide a little bit of direct input to that first revision set of the analysis and modern reports. But the majority of it is really what we look at as confidence building and will give us additional input to rev. those reports to go from Rev. 0 to Rev. 1, and develop another suite of revisions that are upgraded, enhanced, some additional confidence building.

So what you see here that in my--the results that I'm talking about are ones that are really what we look at as in the confidence building framework for site recommendation, with some direct input.

Let me say that talking about an integrated testing and analysis program is a challenge in a way, because what we're doing as we move through the phases of site characterization, as I'm sure the Board is well aware, is we're focusing in on the uncertainties that really seem to matter to total system performance. We're focusing in on those areas where if we're going to try to bound that uncertainty rather than do a full characterization of the uncertainty, we have to have a strong basis for that.

So we're in a situation where we're trying to focus in and do that work which is most critical, necessary and sufficient, is a big challenge because certainly there's some additional work that you need to do in order to make sure your overall representation is good. And so you're balancing between kind of that broader characterization of the site to make sure your processes are understood, and filling in those data gaps where from a performance assessment perspective, we see the highest sensitivity. But that's always a balancing act that we're doing.

The objectives then that we're going to talk about is how we use the next 18 months or so of testing to build confidence in the technical basis, as I just said. We need defensible process models to give us the basis for our total system performance assessment, and as I just said in general terms, in some case, you heard Bob Andrews talk about some of those will be what we call reasonable representation. Some will go to a bounded representation because we believe the uncertainties are such that it's really appropriate to bound it rather than attempt to fully characterize the uncertainties and so with the more reasonable or broader representation.

We also have to make sure that every alternative interpretations that are consistent with the level of information that we have are considered. And as I've pointed out, characterizing the uncertainties to support the sensitivity studies is just absolutely critical. You remember I'm sure some of you are familiar with our peer review panel, gave us a lot of input about this, and said until you convince us you have defensible process models, we're not certain that we can believe your sensitivities and we're not certain that you should. So this is really the focus of the next phase of our testing program.

You saw this chart yesterday in Mike Voegele's talk and I think a couple of other talks. We have now in the revised repository safety strategy that's in DOE review, come up with an enhanced set of factors, and from those, we have a preliminary set of what we're calling principal factors, and Mike Voegele talked you through those yesterday.

The objective here is to get at those particular elements of the system that give us the highest sensitivity to performance, and those are the things we're calling the principal factors.

I think if you look at these, and you look at, as Mike mentioned, the attributes of the system are essentially the same attributes that were in Rev. 0 and Rev. 1 and Rev. 2 of the strategy. So our fundamental system concept hasn't really changed. But what is important is this principal factor, performance of the drip shield, since with the moving forward to EDA II, the new design, we have a drip shield now, so we have to look at all the elements and all of the ways that that impacts our modelling of the system, gives us a different setting for our waste package. So certainly some of what I talk about, and you heard a little bit yesterday, is what does that drip shield do to the environments on the waste package. You know, that gives us a different setting that we have to characterize that we were not really working on prior to adopting EDA II.

Solubility limits of dissolved radionuclides is certainly something that has been a key uncertainty and something that has been looked at in the past, not a new addition, retardation in both the UZ and the SZ, and dilution at the well head. So if you look at all of these, I think the only one that you should recognize as causing us to really look at our test program and make sure that we have the right new efforts ongoing is the performance of the drip shield, and the impact of that on the waste package environment.

Okay, what we're going to do now for the rest of this talk is to simply talk through, picking up where Mark Peters left off, first the testing that's going on for the natural system, and then we'll go to waste package, waste form, materials work that supports the drip shield, as well, and then the engineered barrier system as the overall design concept stands right now.

The way I've set this talk up, in the back of Bob Andrews' talk yesterday, there were some slides that described the kinds of enhancements and improvements he expects to make, or he expects to have in the underlying process models that support the TSPA for SR. And so what I've tried to do is pick up on a few of those just to give you an impression of what the testing and analyses bases will be for some of those improvements that Bob shows will be made in the SR, TSPA process.

So in terms of seepage into drifts, one of the principal factors in our proposed set, one of the things that we're doing here is to give additional bases, and certainly Bo Bodvarsson will talk a little bit about this later, we have some approaches of contrasting the results that you get when you calibrate with test data from both the SF and cross-drift, our two approaches, our continuum modelling, 3-D dual continuum modelling versus discrete fracture modelling.

When you run both of those models and get essentially the same results using the test data that we have, you then have some confidence, number one, that using that continuum modelling approach, which is a much easier approach, is a valid approach, gives you confidence. Also just the fact that you're using two different approaching getting approximately the same result gives you some confidence that you have that process adequately modelled.

So this area is one, seepage into drifts, where in the next 18 months, I think we believe we'll get some additional confidence that will give us a better chance of defending our position at the time of site recommendation with some of the results that I'm going to mention in the rest of the talk.

The unsaturated zone flow and transport, we have some additional realistic 3-D flow fields by using more calibrations. We are getting some lab and field studies that give us better results for the vitric Calico Hills--and this was a big topic yesterday, and I'm sure we'll come back to that today.

The point here is that our lab studies show that we are getting good capillary flow in the vitric Calico Hills. We can show you, or show the community that we need to convince, that the vitric Calico Hills is available for us under the emplacement area, such that we can take credit for sorption in that unit. That will give us a big potential impact on performance.

Conservative estimates for matrix diffusion in the zeolitic Calico Hills, another place where we're getting some additional information that will give us improved basis for the way we model UZ flow and transport, calibrating again with test results from Busted Butte, as I just said.

Okay, for saturated zone flow and transport, again, we have more realistic 3-D flow fields, updated hydrogeologic framework model, and using new geologic mapping results, getting conservative estimates for sorption and matrix diffusion in the alluvium and volcanic aquifers, and we'll come back to this in a little bit as to what information we'll have, kind of in what time frame, using calibration with test data from the C-wells as well as the cooperative program with Nye County that you all heard about in your last meeting.

Okay, what we're going to talk about in the next couple of slides is some of the testing both that continues in the ESF main drift, as well as some of the testing that we intend to do in FY00 and some of it goes into 01 that will give us some additional information from the cross-drift down in that lower lithophysal unit that we haven't really adequately characterized at this point. So this information will give us some really good confirmation that the models, the process models that we're using are adequate, based on the data that we've collected up here in the ESF.

And some of what I'm going to talk about picks up on what Mark Peters had said. Some of what you see on the cross-drift of course is planned, not already in existence, where the alcoves and niches that you see in the main drift for the most part are, I guess all of those are complete. This is a little confusing because it mixes what already exists with what is planned.

For the cross-drift then, the bulkhead studies that Mark talked about yesterday will continue. We'll get useful information on moisture and seepage from the lower lithophysal unit, as well as the lower non-lithophysal unit. Mark showed you along the cross-drift where those units are exposed. Mainly the important information we're getting here on the lower lithophysal gives us a chance to get some additional information there, and some new information there that tells us how representative the results are that we have been getting from the ESF. Similarly in the lower non-lithophysal units, and the Solitario Canyon Fault zone.

For the cross-drift and niche studies that crossover Alcove 8, at the crossover alcove here is where we're talking about--we'll have flow and seepage testing going on between the cross-drift and Niche 3 in the ESF, so this will give us some really valuable information, providing field scale data for the important UZ flow seepage and matrix diffusion. But the important point here is by setting that test up the way it's designed--I'll have a picture in a minute that will help understand and visualize that test--we are going to be able to get seepage and matrix diffusion measurements over scales of tens of meters. You know, most of the measurements so far have been on the order of a meter, or so. This will get us out into tens of meters that begins to get at the scale where it's really important to look at for repository performance.

Okay, in Niche 5, also along the cross-drift, we do some hydrologic characterization with the air permeability and seepage testing in some systematic boreholes, and this again will get at seepage process data, data on variability and hydrologic parameters, and again get at improving the overall seepage model in that lower lithophysal unit, which makes up such a large percentage of the repository host rock.

Okay, a picture now for the cross-over alcove, the one at the intersection or at the point where the main drift is crossed over by the cross-drift. This is the Alcove 8 setup. This is the one that will allow us to get at some tens of meters of scale of seepage and infiltration. This will be a really valuable test.

And on this one now, I think this one I have coming up in just a minute, some dates that will tell you what our current plans are, given budget assumptions, for when we should start getting some test results from this one, as well as from the next one, because I know that that's of interest.

For Niche 5, Niche 5 is out here almost under the crest. For Niche 5 again, the kind of testing we could do to get at the performance of the lower lithophysal unit, very important testing, and the question of schedule--I think this one is probably not as easy to talk to as the next one, but you'll notice that what we've highlighted is that for, this one is Alcove 8, which is the crossover testing, Niche 5 out in the middle of the cross-drift, and then the systematic characterization in the boreholes, this would be all of these feeding to Rev. 1, meaning in the time frame of July of 00.

So we're at the point where we can get some information that will help us to build confidence in what we had in Rev. 0, as we do Rev. 1, begin to gain confidence that we have the right set of processes, particularly in this lower lithophysal unit that I know the Board had some concern about.

The next page I think gives you a better picture of that schedule. In terms of Alcove 8, the current plan is to start very soon with the excavation, starting with the drill and blasting, and then roadheader. Coring to start in January. Testing setup in February. And you saw when the first feed of data comes from Alcove 8 on the previous network chart.

For Niche 5, again, starting early in calendar year 00 with the testing setup, the second phase coming in the middle of 00, and the systematic characterization holes out in the April and May time frame.

So I think you can see that we are putting some high priority on getting some data from the cross-drift as soon as reasonably possible, to get at this question of representativeness of ESF results when they do not represent that lower lithophysal unit.

Okay, now, talking about ESF results, the additional work that will continue in ESF, we talk about Alcove 1 and we'll talk about 7, and then the niche studies also. Okay, for the Alcove 1 and niche studies, this picks up on what Mark talked about in terms of flow and seepage testing that helps us with the El Nino effects. One important thing that we can do with the niche studies that's planned and isn't quite described on this slide completely, but one of the things we want to get at is the variability that will help us to understand, and Bo will certainly elaborate on this, this whole question of whether we have a seepage threshold in effect. And through the niche studies that we have set up for FY00, we are going to be able to move from one that's completed in a Niche 2 that has a medium permeability setting, to Niche 3 which is going on right now in a low permeability setting, to Niche 4 with high permeability in 00.

So what we should be able to do there is to get a sense at least for how that seepage threshold performs in rocks of different permeability, and that should give some important information to us in order to determine whether we are going to be able to use the seepage threshold as an actual performance constraint.

So the overall testing then improves the confidence in seepage and matrix diffusion, expanded basis for climate effects because we're looking at the variability in infiltration rates and the impact that has on seepage.

Alcove 7 moisture monitoring, this is the one that Mark talked about yesterday where very interestingly, we see the return in that area that has been bulkheaded off around the Ghost Dance Fault, you see it returning to ambient conditions even though the fault is present. So that's giving you some good information. If that continues to show, that is, if that continues to be observed, then we certainly have some good indication of what role at least that the current conditions of Ghost Dance Fault is playing or not playing.

For the validation studies relative to the chlorine tracers, chlorine and chloride mass balance, there is, as I think Mark mentioned this yesterday, there are two ESF bomb pulse locations, Sundance Fault and Drillhole Wash Fault zones, where we will do some additional sampling and measurement to increase the understanding of whether these are in fact zones where we have preferential pathways, also using the chloride distribution to calibrate UZ flow and transport, which Bo will come back to later, and completing some mass balance studies. So this whole area is one that is in progress, will continue to benefit from our understanding of that work as we move forward from current understanding into Rev. 1.

For Busted Butte, again, it's just a continuation of the data analysis, but going to that Phase II study that Mark showed you the picture where it's a much larger volume of rock that's being characterized, gives us the important matrix diffusion and sorption data in the non-welded Calico Hills, and we know we have an issue there that we've talked with you about how representative or how applicable that is to the volume of rock under the emplacement area, and that is something that we are going to have to spend some time considering how we make that case.

And I think the important thing to understand, given the discussion we had yesterday, is that exactly how the vitric and zeolitic areas are displayed or aligned isn't really the important factor. The important factor is what kind of reliance we're going to place on those two types of units within the Calico in the performance assessment. You know, what are we going to try to defend, in my view at least, not exactly where the transitions are in the rock properties.

For testing and analysis addressing thermal effects, the thermal test continues of course for four years, cool down for four years, and post-test characterization. You all know, you've had many briefings on this test, large scale thermal effects on seepage, helping us to get bounds on chemistry and the amount of water contacting the EBS and the waste package, and we'll look at this test in terms of ways that it can help us address the questions related to the lower thermal loads.

You heard Mark yesterday mention that we are seeing some moisture changes even below the boiling temperature zone, and that that's important to understand what kind of thermal effects will you have, even if you don't boil. You know, if you go to the longer term ventilation period, you end up with a non-boiling drift wall, you're still going to have to look at what kinds of effects you have because of the elevated temperature.

Cross drift thermal test is planned to get that same kind of information in the lower lithophysal, which you know as I mentioned is the majority of the host rock. That will expand our data for thermal effects on seepage, performance of the drip shield, giving us a basis for performance of our drip shield and waste packages, give us increased confidence in the process models. And this one is out in license application time frame under current schedules. This one certainly isn't going to be set up and giving us any results that are going to be useful to us in site recommendation time frame under current schedules.

The saturated zone principal factor, important collaboration going on here with the Nye County program that you've heard about. The role of the alluvial aquifer has certainly become something of interest to us. We won't be able to get information on that, particularly in the early site recommendation time frame, but we certainly will get some additional information to help us with flow path characterization and some at least hints of what kind of performance you might get out of the alluvial aquifer.

Interactions between tuff and carbonate aquifers are important, as well as the field scale transport in the saturated zone.

Now, natural analogs came up several times yesterday, and the Pena Blanca site is one that we have talked about I think with you, and I'll mention a couple points about that, and then there are other analog sites that will be looked at. There's a little bit of work funded in 00 that will help us I think bring natural analogs in to the extent that we could use them to help validate models.

Pena Blanca analog site for transport of uranium and daughter products, the past work has focused on the open versus closed system behavior, timing and rate of migration of the uranium and thorium type of isotopes. The results so far suggest stability of these isotopes over long time frames, on the order of 300,000 years. So you're talking about some useful information, perhaps not as useful for our site as it could be, but it's still interesting, and from the standpoint of building confidence in the general way that these elements behave in a natural setting, it is probably of use to us.

There will be some planned drilling to provide rock and water samples that will give us some initial validation of transport rates.

The other analogs, and I'll just mention these, and I think Bo will pick up on a couple of these, both INEEL and Hanford, we have some work in our FY00 plans to look at, particularly at Hanford, at tritium plume migration in saturated zone alluvium. That should help us build some confidence in handling dispersion. We can compare results of our modelling with the PNL results of the modelling that they're doing for that plume. So that's at least one area where we can do a little bit of benchmarking and/or building confidence, similar some plume modelling at INEEL, which I'll leave for Bo to talk about.

Another one that is interesting, I think Walter Matyskiela mentioned yesterday about potential for using any kind of geothermal or igneous intrusion as a potential analog for mineral alteration. We have a little field study planned at NTS to look at a cell to see whether or not you can get any kind of an understanding of potential alteration by looking at igneous intrusive bodies, or geothermal settings.

I think this is one, just as an aside, we've looked at this a number of times, but one of the things that the geochemists have often claimed is that you have, in a sense at least, a nice natural analog right in Yucca Mountain because you know the volcanic rocks there have come through that temperature alteration period as they were erupted and cooled. And so when you kind of go backwards and look at the kinds of alterations that have occurred, you in a sense can gain a lot of understanding about the kind of alteration you will have when you heat them back up.

Other ones you might have heard about, other analog studies you've heard about that are not being worked on in 00 are--there was some work at a Russian site, as well as Okro that we have talked about in the past, so we're not doing any work on those in 00. And right now, nothing is planned with regard to anything at the Nevada Test Site outside of our work.

Okay, moving along to the waste package and waste form, including the materials testing that supports drip shield, since we kind of lumped the materials testing together because it makes the most sense, since both titanium and Alloy-22 need to be looked at through the same set of conditions and environments. The improvements that we think we will see, and you see this if you look in Bob's backup yesterday, performance of the waste package, we're going to have a better analytical basis, mechanistic analysis for the kinds of defects, the kinds of early failures that we will need to include in our modelling.

We're going to include additional corrosion mechanisms, stress corrosion cracking, get additional confidence of long term phase stability, and then the effects of aging, thermal aging particularly, and I think Joe Farmer will have more to say about these when he talks about validation of these models this afternoon.

Also, of course, new data on corrosion rates, and as Bob mentioned yesterday, we are moving from a bases pretty much from our expert elicitation panel inputs, now to having some good laboratory data, as well as some data that we can bring in from other industrial experience in the case of titanium that will give us some additional confidence in our modelling.

Other improvements; the solubility limits for dissolved radionuclides. Here's one where a reasonably bounded representation for SR will be our basis. There is new data on the relatively immobile radionuclides. We'll talk about these a little bit more as I go through what the test programs actually are.

There's some related factors, not principal factors, as we have them characterized now that will also be improved. You know, you understand that as we walk this line between principal factors and other factors, one of the key points is that we have to have enough understanding and enough bases for the ones that we are not calling principal factors to be able to convince the world that we have that right, that in fact they are not major contributors to performance, and they don't have major sensitivity if we go to a bounding representation for that factor.

So colloid-associated radionuclide concentrations is certainly one of those, and I know colloids came up several times yesterday, both in near field as well as far field.

We'll have an improved colloid formation model, some new data on sorption/desorption, and the Americium colloid data will be added. There's a question, I think in my notes I had a question that I didn't get a chance to follow up on. I don't think that will be into the Rev. 0 type or Rev. 1 type time frame. I think that's a little bit further out.

Cladding degradation model, direct evaluation of clad unzipping, we have some experimental work going on at Argonne that will give us some direct laboratory data on this. Conservative bounds on initial defects, we'll talk about that a little bit more in a minute.

Okay, this one is just to give us a chance to look at a picture. I think you've seen the current concept. Mark Peters had a couple of figures I think that show you essentially a corrugated drip shield over the new waste package design with the Alloy-22 on the outside. The drip shield concept is being looked at. Certainly we're not locked into this yet, but there are some questions about the way that type of drip shield will perform.

As I mentioned earlier, one of the key things that this has done for us is to cause us to ask the question what kind of environment will exist below that drip shield on the surface of the waste package. And so in terms of new drivers for testing, that's one that is really important to us.

I think I might mention on that one one other point. One of the questions, or another issue that's been raised is how important it is to look at the supporting mechanism, the pallet or whatever type of support we finally end up using, under the waste package and the relationship between the waste package, that pallet and the invert, and even the invert materials, some kind of a ballast.

The question of whether you have problems at those contacts, and the exact type of material you should use is one that is currently being evaluated. Further optimization will certainly occur there.

Okay, the elements that are most important to performance, this came up yesterday, I think Paul Craig asked a question about how we will get at any kind of fabrication, any kind of testing techniques that will help you reduce the probability of early failure. The issue of how you're going to reduce any kind of stresses that occur in your welding, at the welded units, we know that's going to be a big issue, and I think Joe Farmer and I spoke with Paul Craig about that question yesterday. Livermore has some approaches that they're looking at to reduce the stresses such that the welds will not be a preferential point of corrosion. We think we'll have a sound basis for our assumptions for early failure in the site recommendation time phase.

The kinds of techniques that we're going to use for non-destructive testing are standard approaches, proven technology, ultrasonics that are used by the nuclear industry, so we don't think that we're going to have a major technology problem there in terms of being able to test the condition of those welds.

Upgrading the process model with additional degradation modes, as I mentioned, that's one thing that TSP expects from the waste package area. Localized and general corrosion tests are in progress at a range of concentrations. General corrosion rates are very low, and you've heard some discussions of these, and we'll hear further from Joe Farmer. Pitting corrosion has been demonstrated not to be a significant factor, we believe, but there is additional testing underway that will help us build confidence in that position.

Okay, we have improved data for stress corrosion cracking for the Alloy 22, for Titanium 7 and the stainless steel now that's being used as our structural material inside of the Alloy 22.

Industry experience and test results on stress corrosion cracking and crack growth under repository relevant conditions are available. This is one where I think Dr. Sagüés yesterday had indicated that he felt that we had a fairly limited amount of information available on Titanium 7. Our folks have spent a lot of time going out and gathering what information there is, and we have a draft analysis and modelling report available now that is in review that pulls those nuclear and non-nuclear industry experiences together and does get the information available on stress corrosion cracking, crevice corrosion, hydrogen embrittlement in one place. And we actually feel that there is a fair bit of information available on Titanium 7. So our view is it isn't quite as bleak as what you claimed it was yesterday, but that certainly is available for review at some point, and you can draw your conclusion about what we've pulled together. We think that data will be adequate to benchmark the model and determine susceptibility to these modes by site recommendation time frame.

Another issue that is of importance is the long term phase stability and thermal aging. Here, the issue is the potential for precipitation of intermetallic phases that cause areas that are more susceptible to corrosion or the hydrogen embrittlement problem that Titanium shows, and stress corrosion susceptibility.

Here, we have some accelerated testing going on. The hydrogen induced cracking concern, there are some notch specimens that are being run under bounding conditions, so these are accelerated, extreme type of tests just to get some information that will give us some early indication of whether these are issues.

You know, the general corrosion community attitude seems to be that they are not, but we understand that we have to have some level of test data available to give us some basis for taking the position that the probability of those kinds of changes causing problems is low.

Okay, again, another area that's of concern is stability of the passive corrosion films on Alloy 22 and Titanium 7. We have some information now being pulled together, again from a lot of different sources, and one of the things I think you'll find is that from both this Board, as well as from our peer review panel, they have in the past told us we haven't been creative about going out and bringing in information from outside of the project, information from nuclear or non-nuclear sources that is relevant and can be helpful to us, and I think you will see our people have done a lot of that as we moved into this phase of the program, trying to document the basis for some of our judgment that has been challenged.

Stability of both Alloy 22 and Titanium grades that are not too unlike Titanium 7 have been demonstrated after a year of exposure, and I think Joe will talk about those testing results from Livermore. Alloy C, which is rather similar to Alloy 22, an example from a nice nature analog where it's been exposed for 60 years in a marine environment, and that one shows basically original condition. Still has its shiny surface.

Another natural analog, a type of nickel/iron mineral exposed in stream beds shows no film breakdown. So we're looking for every kind of source we can, with the big question recognized to be how do you take the laboratory data of a few years, months and years, and extrapolate the long time frames. We know that will be the big challenge. And then some additional testing, again that Joe can talk about, where we're looking at corrosion under oxide deposits on the waste package.

You know, one of the issues here that I didn't mention is when you have the drip shield in place, the environment on the surface of the waste package is different, and the question and one of the challenges is is that environment going to be pristine, where you have basically very clean surface and where you have absence of salt deposits. What we have to look at, and that's one of the things I'll talk about in the drip shield test, is what kind of a chemical environment will you create under that drip shield on the surface of the waste package, because that will be really key to the performance of the waste package in our new design concept.

Okay, the surface environment. Some new data indicate boiling points and pH can be higher than previously assumed. I think you heard this in the previous meeting. 115 to 125 degrees C boiling point. Phs can go high. On the other hand, if you have some of the other effects driving you to lower pHs, the question is what will that environment look like through time and space.

Experimental modelling effort will provide expected range of environments, and the models will be benchmarked, uncertainties bounded for SR.

Okay, on the solubility side of radionuclides, plutonium, uranium and neptunium, some of those key solubilities are being re-evaluated and we'll bound those in our models for SR.

Colloidal radionuclides, again potential mechanism for transport, and those will go toward the bounded uncertainty for site recommendation.

Cladding performance is one where we are getting some additional information, bounded uncertainties for the models for SR, but the initial state will be defined better than we had for viability assessment, with the fraction breached at receipt, the degradation rates, meaning the fraction breached with time, and the unzipping rate, surface area for dissolution and transport resistance, with some additional tests that are going on, as I mentioned, at Argonne.

Waste form degradation rates, bounding rates will be used for site recommendation. And some of these are not much of a change from what we did for viability assessment.

But talking about engineered barrier system, the improvements that you see in what Bob presented, new drip shield degradation model, we'll have a mechanistic analysis of manufacturing defects. As I mentioned, that's being done for both materials, both the Titanium and the Alloy 22. We'll include the hydrogen induced cracking, but our design is set up to isolate the Titanium from hydrogen sources, so there won't be a direct source of hydrogen from carbon steel, or from anything that could give the Titanium a potential for hydrogen induced embrittlement.

And of course our overall performance of the drip shield, one of the things we have to look at is what kind of a rock fall, you know, assuming that you have backfill over the drip shield, the rock fall should not be a big issue. The drip shield should be protected by the backfill. But the question of rock fall, as well as seismic loading have to be looked at, because one of the concerns is with the type of overlap that we have in the current drip shield design, is if you have some seismic shaking, will you get some separation, some gaps developing, and if you have backfill sitting on there, will the backfill trickle down between the gaps that develop in your drip shield.

So this area is one that is really receiving intensive thought and study, and is one that is new to us and, therefore, the models that we have to develop are relatively new and will be moved on to the maximum extent we can as a basis for the TSPA analyses for SR.

This just gives you a sense from the engineered barrier system perspective of the various parts of the system that have to be looked at. Clearly, it's important to us, and I think yesterday, someone mentioned, you know, what is the real purpose of the drip shield testing that's going on, and it's very important to get at where the water goes, water distribution, if it's diverted, where it's diverted to, where the drainage occurs, what the thermohydrologic chemical conditions are in that area under the drip shield.

Physical, chemical environmental model, the transport model, once you get anything released, how the material moves through the invert. And then there's a number of other sub-models that are pieces of this that all go together to give you the abstraction. And, of course, coming in from the waste package side, or the materials side, is the degradation performance of the EBS.

So putting together this overall model for the EBS, for the drip shield and the relationship with the waste package is really a major focus of the work in the next 18 months.

Okay, the performance of the drip shield clearly depends on where the water goes, how the water is excluded. The backfill drip shield flow processes are critical. Thermal effects on that flow, any kind of impact of the thermal effects on the EBS materials is critical. And, as I mentioned, the degradation modes, any kind of shifting, if you have an overlap, any potential failure at those gaps or cracks.

We have pilot scale testing and a column test that I'll mention going on to get at this information. Water distribution and removal model is being developed, and Mark mentioned that yesterday and showed you some pictures of the kinds of testing that is set up and in fact started right now. The in-drift thermohydrologic chemical changes in EBS materials are also being looked at in that testing.

And then finally, this was also mentioned yesterday, seepage into the drifts is affected by their geometry, and part of the work in this area is to get a good drift degradation model in place that considers frequency of rock fall, block sizes, total extent, timing, because we understand the importance of the geometry on the seepage.

There are a number of early component testing that have been completed in this facility at what we call the Atlas Facility, and all of these give us a good bases for designing the next phases of the EBS of the drip shield testing. We had the pilot scale test, and I think some of you have visited that facility, for the Richard's Barrier, which was very effective. It did divert water as we predicted it would. Some pilot scale testing of single backfills, some flow visualization tests to look at the Richard's Barrier in a fairly simplistic manner, some other laboratory tests to get at diffusion eoefficients for the different options for backfill, as well as invert material.

So these results are really there and are available to be used in building our Rev. 0 bases for the site recommendation.

For the EBS testing and analysis as we move out, we've got pilot scale test Number 4, which is a drip shield with backfill. This backfill is a fine backfill. This is different than the next one I'll mention, which has a coarser backfill. The purpose of this one will be to validate models of moisture and chemical responses for our EDA II configuration and verify the conditions that control condensation under the drip shield.

As I mentioned, the real concern here is what kind of environment do you create by putting this drip shield in place. There are some who have challenged us and said are you sure that the complexities that you're adding by putting this drip shield over your waste package is worth the benefit you're getting. So we are going to have to be able to answer that question.

The test design for this drip shield pilot scale Test 4, sand, fine sand as a backfill, crushed tuff invert. I might mention on the case of the invert, there's questions being looked at in terms of what would be the best material, whether crushed tuff is the best material is still open for discussion. Scale model drip shield, and simulated waste package will be at 80 degrees C. Drift wall will be kept at 60 degrees C. in a manner that Mark showed you yesterday in the configuration of the test. The inflow rate will be varied to relate seepage with the kinds of conditions you see in this experiment.

One additional on that one is that there's some interesting thought that perhaps because we saw the Richard's Barrier perform so well, there's some thought that the contrast and permeability between the backfill sitting on top of the drip shield, that you might actually get a Richard's Barrier type of performance barrier there, such that the water won't actually move from the backfill onto the surface of the dripshield, that it will be diverted and move through the backfill. And that's one of the things that we really want to look at in this test.

Pilot Scale Test 5, big changes that go to the coarse backfill. Verify the conditions that control condensation, and again look at the models for moisture and chemical response, but with a much coarser backfill, similar conditions for the rest of it. So this will give us a chance to look at the variability in conditions that is caused by a change in the nature of the backfill.

The saturated alteration test is interesting. One of the things that has become a concern with the current design is what happens if you plug either the backfill or the invert material such that you create some ponding and your waste package at some point in time in the future has dropped down and it's sitting in these little ponds of water. And so the question has become have you created another failure mode, or a new failure mode that you really have to show will not be a problem, or if it is, maybe that becomes the most likely failure mode, is this dropping of the waste package into the invert.

So this experiment is set up to cause--it's a column test and it's set up to actually cause some accelerated build-up of salts, take J-13 water and reflux it in through the crushed tuff type of material, and see what kinds of salts develop as you vent the vapor and accumulate the salts and minerals. So do something in such a manner that you can quickly see if this invert plugging and potential for ponding is really an issue.

Calibrate the thermohydrologic chemical models to whatever alteration you see, also do some of the same kind of testing, but in an unsaturated column test.

Finally, testing has been expanded to include new and revised SR design, improved waste package, backfill, drip shield. We've talked about testing and analysis program is designed to focus on improvements to the key process models and to focus in on the principal factors that are correlated with those key process models, provide a sound technical basis for reasonable representations where that's appropriate, for bounded where necessary, and alternative models, basis for considering alternative models where that's appropriate, and also define the uncertainties so we can support sensitivity studies.

So this hopefully gives you a picture of that next phase between now and the time that the site recommendation formally goes out. A lot of additional work, a lot of additional information should become available to help us build confidence that the way we've represented the system in Rev. 0 reports is adequate and appropriate. Thank you.

PARIZEK: Thank you, Jean. Any questions from the

Board? Debra?

KNOPMAN: Knopman, Board. Jean, this is quite a list of activities, and I appreciate that you went through all this with us.

My question concerns a discussion you started in on about the added complexity that a drip shield brings, and you had I guess it was--you had a slide that had a pretty detailed list of the different, Slide 32, on all the different aspects of the drip shield that you're going to need to be looking at.

Have you gotten to the point where this work is--it's not just a question of prioritized, but put into some kind of critical path, framework, so that you would know sooner rather than later whether this is really worth the added complexity? That is, do you end up creating more problems and more uncertainty for yourselves than you would if you, instead, took the money and resources that will go to this and put it into other aspects of the system? I don't know the answer to the question. I'm just wondering if you've kind of set this up in a way that you'll know whether you cross some threshold or not soon rather than 18 months from now, and the thing has just not come together.

YOUNKER: It is really a good question, and I think we probably need to look at the way we have the EBS drip shield test phase, and look and see whether there are some points in time when we should ask ourselves that question, pull all the information together and have a hard look at how good is that pre-test and post-test modelling, you know, how good are the results relative to what we have been able to establish, and determine whether we're getting a handle on, you know, what kind of an environment are we creating, how much reflux or how much condensation and salt deposit are we really seeing. It's a very good point.

KNOPMAN: There are also I would think two kinds of risk situations you'd want to consider. One is sort of the what might be considered normal conditions of just wear and tear, versus the low probability, high impact type events where some kind of shaking motion would topple the drip shield, and what you have is a bunch of rubble, and none of your modelling will have been able to do anything with backfill and rubble of a drip shield sitting on top of the waste package. But there's some probability associated with that kind of outcome.

So we'll look forward to seeing more analysis from your end on how you're going to proceed here, because that would certainly be a concern of mine, that you're going to put a lot of effort, kind of go off on all these different directions, and not have a clear decision making framework.

YOUNKER: Yeah, I think the designers are fairly confident that they can build a drip shield, build and install a drip shield that will withstand the kind of seismic shaking and the kind of design basis rock falls that we anticipate. So I think that side of it, my impression is is probably less of a challenge than getting at the way the water will move and what kind of environment we'll create on the surface of the waste package by having that drip shield in place.

You know, initially I think that I know Dr. Bullen had discussions with us about this where there have been initially some claims that, gee, it was going to be a pretty pristine environment, but then you think about the fact that you've installed backfill, and certainly it would be hard to keep a dust free environment while you're installing backfill. So you know there's going to be some dust. You know there's going to be some reflux of water during the time that you're in the thermal phase, and you know there will be some evaporation and precipitation, some salt build-up.

So I think we're really getting focused in on the questions we need to answer, but we're certainly not at the point of having definitive points in time to find where we take a critical look and make some decisions about whether, you know, the trade-off is going the right direction.

PARIZEK: Priscilla Nelson?

NELSON: Maybe these questions, at least one of them, should be deferred for Joe Farmer, but they're little questions.

First of all, on the ECRB Alcove 8 to ESF Niche 3 test, as I recall, the stratigraphy is such that both the lith and the nonlith are involved in that flow path.

YOUNKER: I think that's right.

NELSON: Is there going to be an attempt or instrumentation to separate out the performance of the two different rock units in that flow path?

YOUNKER: I'm not familiar enough with the detailed design--is Mark Peters Here?

NELSON: Is Mark still here?

YOUNKER: Mark, did you catch Priscilla's question?

PETERS: Mark Peters, M&O. You're right. It's about--it starts in the upper lith. It's about 18 meters to Niche 3 below. So it's roughly two-thirds upper lith, one-third middle nonlith.

If you remember the picture, there's boreholes coming from up and below. So they'll be instrumented in both units, so we should be able to pick up some of the changes in flow paths as we go between the different units.

NELSON: Yeah, you might be able to. I'm wondering even whether there might be some other excavation that would actually remove it and get one rock unit at one point later. Anyway, that's fine.

YOUNKER: Priscilla, we'll pull up the picture just so what Mark said makes sense. We're almost there.

PETERS: There's the unit.

PRISCILLA: The bottom third is in the--

PETERS: Right. So those red boreholes actually penetrate up into the upper lith, and the upper boreholes penetrate down into the middle non.

PRISCILLA: Right. But the particular attention to try to separate out the performance of the two units is only going to be done through borehole measurements?

PETERS: Correct.

PRISCILLA: Okay. Stay there just for a second, because you brought up Busted Butte, and I guess we had some discussions yesterday about it and they had to do with the vitric and the non-vitric portions and where the zeolites were. And a lot of that discussion has always focused on the matrix or the petrographic characteristics of the intact rock, and how similar they were from one place to the other. And so is the testing of Busted Butte really focused towards matrix activity?

I'm wondering do you know anything about the fracture frequency information for these units? I mean, with vertical boreholes, you don't get very much information on fracture frequency information, and the importance of fracture flow in the Calico Hills.

YOUNKER: We do have a table that summarizes our best estimates of the fracture frequencies in a letter that, Mark, you and I put together that describes the expected differences.

PETERS: For the Calico? We're talking Calico Hills here; correct?

YOUNKER: Yeah, I think so.

PETERS: You're correct. The vertical boreholes make it very difficult to get good fracture frequency information in the Calico, so we don't have a tremendous amount of information on that. I think the key is is how you assume it acts in the model. And Bo, I think, will probably address that in some of his talks. We don't have a clear understanding, a real good understanding of the fracture frequency underneath the repository because the boreholes just don't give a lot of that good information. We have some information from outcrops, but not under the repository.

NELSON: Thank you.

PARIZEK: Dan Bullen?

BULLEN: Bullen, Board. Jean, you mentioned the effects of the addition of the drip shield on the waste package environment. But one of the big significant changes that's been made since VA is the fact that you've changed the waste package design such that the wall is thinner, so the radiation field is up a couple orders of magnitude. And what I don't see, or what I'd like to see, I guess, are issues addressed with respect to the effect of the radiation environment on the degradation of the drip shield underneath in that area where the radiolysis, you know, may have a significant effect on drip shield performance.

And so do you have a plan, or are there scale tests or tests that might be addressing that?

YOUNKER: Yeah, I think Joe is going to talk about it a little bit later.

FARMER: In regard to the gamma radiolysis, you know, early in the mid Eighties, we did the gamma pit studies with 300 series stainless steel, and we've been wanting to restart those efforts but haven't been able to. So in lieu of doing gamma pit studies, we've now done studies where we've purposely added hydrogen peroxide at various levels and looked at the impacts of the hydrogen peroxide on the corrosion potential and the threshold potentials of the corrosion resistant materials such as Alloy 22, Titanium grade 7, et cetera, and we'll show you at least one or two examples of that this afternoon.

BULLEN: Joe, before you leave, those potentials are actually the addition of hydrogen peroxide to a water environment; right? Not to a thin film?

FARMER: That's correct. Actually, what we have done is we have standardized all of our test media. As you know from the long-term corrosion test facility, we have simulated dilute water, concentrated water, acidified water, so on and so forth. We've now added to those generic test media some new environments which are basically more or less fully saturated. To those generic standardized test media that we're using across the project at Livermore, at General Electric and various institutes that are working on the project, we add hydrogen peroxide. And it's more or less like a titration experiment, you know, we'll add hydrogen peroxide basically increasing the hydrogen peroxide concentration at eight part per million steps, up to the point where you no longer see any increase in corrosion potential.

And, of course, the important issue is to make sure that you don't push the open circuit corrosion potential above any threshold for localized attack. And as you'll see this afternoon, that is in fact the case. You can add as much hydrogen peroxide as plausible, and even go beyond that, and you can't push the corrosion potential for these corrosion resistant materials into a regime where we would expect any sort of destabilization of the passive film. And of course that isn't the case with 300 series stainless steels, and that's the reason we picked Alloy 22 over the 300 series stainless steels.

BULLEN: With respect to the Titanium that you're testing, Joe, are you doing the same kind of tests for Grade

7? And actually, the other question I had was that as you standardize your tests and add the titration of the hydrogen peroxide, does it end up in the vapor phase of those tests or not?

FARMER: We have not done vapor phase experiments with the hydrogen peroxide yet. That's probably something that Greg Gdowski would, you know, ultimately do in one of his experimental apparatus. But we haven't done the vapor phase hydrogen peroxide experiment yet. For a lot of these fast track experiments, we're having to use some of the tried and true techniques like cyclic polarization.

BULLEN: I understand that. But I just think that sort of along the lines before you actually commit yourself to making a Titanium Grade 7 drip shield, you ought to take a look at the fact that the vapor phase above the waste package is going to be one of the key issues.

But, thank you. We'll talk about this this afternoon, and I'll defer. I have one more question for Jean.

In the Atlas facility test that you identified, you had Series II and IV and V, and you basically have a test that's high temperature with respect to a waste package or a surrogate waste package of 80 degrees and a drip shield or wall temperature of 60 degrees C?

YOUNKER: Right.

BULLEN: Do you expect that to be applicable or directly relevant to a 96 degree C. type of environment, or do you think there will be some changes with respect to that extra 20 degrees that might have a problem?

YOUNKER: I think we're going to have to look at that to make sure that the test is exactly right for the EDA II concept, and since we have--the EDA II concept is a, kind of has operating conditions of either closure at 50 or closure at 125, clearly we're going to have to look at the way that test can be configured to best give us information for either of those. So that's a good point.

BULLEN: Can you scale the Atlas facility to 120 degrees C., or is it not quite--I mean, you get close to pressurization problems there?

YOUNKER: Yeah, I don't think it was set up to do that. Jim, do you want to comment on that? That's the reason why we're constrained by those temperatures.

BLINKER: Jim Blink from the M&O. I think those experiments are designed to give insight rather than to be full prototypical tests. They're at the quarter scale. They wanted to set up a Delta T across the waste package to drift wall, a higher Delta T than we would see in a normal situation, to try to drive the condensation process and see where the water formed and where it dripped and whether it concentrated in the invert in certain ways.

They will apply those results to calibrate models that will then be applied to the range of environments expected as time progresses in the repository.

BULLEN: Thank you. And, Mr. Chairman, I'll defer, because I saw a whole bunch of hands go up, so I'll stop asking questions now.

PARIZEK: Don Runnells?

RUNNELLS: Jean, you didn't mention any of the lab experiments that were going on about a year ago at Los Alamos on retardation, particularly of neptunium. Are those continuing as well?

YOUNKER: Yes, I think that's some of the basis for the improved data that we'll use. I'm not real familiar with those, and I'm not sure if we have anybody here who is.

RUNNELLS: Those were column experiments?

YOUNKER: Column experiments, yes.

RUNNELLS: Under strongly reducing conditions.

YOUNKER: Yes. I think those are still carried into FY 00.

RUNNELLS: Okay. So they're continuing?

I have a question that's just I guess a little bit facetious, but maybe not entirely. The drip shields sound like a lot of Titanium to me. How does the amount of Titanium that's projected to be used in drip shields compare to the world's annual production of Titanium? Do you know for a fact that you can buy that much Titanium at the rate that you need it?

YOUNKER: Yeah, I don't think that's an issue, and I think that has been looked at. They are only 20 millimeters thick, so they aren't exactly--it isn't like as if it's a huge amount. But I don't think that's an issue.

RUNNELLS: Okay.

YOUNKER: Jim was on the team that recommended Titanium be considered.

BLINK: Jim Blink from the M&O. I'm trying to remember from when we discussed this in LADS, and I think it was something like a 3 or 4 per cent of the current demand that would be required per year for a period of several years.

RUNNELLS: Okay. That's reassuring. I had no idea what that figure was.

YOUNKER: I remember we did ask ourselves that question.

RUNNELLS: Okay, that's good. One other question about the drip shields. They do, as Debra said, introduce so much complexity, can you just recap very briefly the history of why they have appeared in the design? At some point, somebody said we need something else. Maybe it's a drip shield. What happened there to cause that?

YOUNKER: If you recall some of the discussions yesterday that Mike Voegele had about when you look at the importance analysis and when you look at the contributions from the natural barriers, which are significant at this site, no question, when you add in the waste package, which we know we're going to use a waste package of some reasonable level of robustness, you look at that and you ask yourself the question from the results of the importance analysis, do you want to have all of your defense resting on that waste package barrier, or do you want to do something to give yourself a second line of defense. And that drip shield really represents that.

It gives you not only protection of your waste package, your primary barrier from water, assuming that we can get at this question of the environmental conditions under the waste package, but it also gives you a second line of defense. And I think that's the primary reason. Having a drip shield there really is an independent, or almost independent barrier that can give you protection for your waste package and gives you that independent confidence that you have an adequate system.

RUNNELLS: Defense-in-Depth?

YOUNKER: Yes.

RUNNELLS: Okay, thank you.

PARIZEK: Jeff Wong?

WONG: Let me struggle with this question. Mike Voegele earlier, or yesterday, said that concluded confidence will not be adequate, unless the natural systems can be demonstrated to contribute significantly. And I look at the timeline that Steve Brocum had in his presentation, and I look at your testing, so I guess I'd ask you what's your definition of increasing confidence? Does that mean decreasing uncertainty in performance? And do all of your tests that you have underway within the timeframe of the SR, how much confidence do you expect to increase by?

YOUNKER: I think that our sense is that at Rev. 0, at the time that we're building--I think yesterday, it was made very clear a couple times that, you know, the fundamental technical basis that we have for TSPA SR is pretty much in place right now. Rev. 0s are being written, many of the Rev. 0s of our analysis and modelling reports are heading into review. And so, you know, that fundamental bases is pretty much there, and as Bob explained, and will explain further, there's an important distinction between what we are able to use as direct input, which is what is in this Rev. 0, and what we will use to build our confidence and further enhance the Rev. 0 as we go to a Rev. 1 phase for the analysis and modelling of course in the process model reports.

So I guess my view is that, you know, my sense is from talking to the scientific and engineering folks that support us, that our confidence is pretty good in that representation that we're going to give Bob, or that Bob is going to make and that we're going to give the process bases for.

As it stands now, you know, we had a viability assessment was a good trial run. We had a lot of criticism of the areas where there are big uncertainties and where there are gaps. We focused this program as much as we could to get at those in a short time frame, with some accelerated testing. You know, some of it won't deliver as much as we would like, but I think someone answered the question this way yesterday, you know, in those areas, if what we do is continue to build confidence and confirm that the approach and the representation we have is pretty good, then I think our confidence will continue to grow as we go through the testing in the next 18 months, and we'll have I think a strong bases for our site recommendation.

If in some areas we get some surprises, we will have to go back and look at it and see what difference it makes. We'll have to look at whether that surprise and that difference down at the process level really matters when you roll it through abstraction and total system performance.

So the whole issue will be how important is that news or that surprise to the fundamental performance of the system.

WONG: Then the seven factors that you've listed, or have been listed in the previous presentations, are those factors that you have low confidence in?

YOUNKER: That we have?

WONG: Low confidence in.

YOUNKER: No, no, not at all. In fact, I think the confidence in both the other factors and the principal factors is highly variable. When you see what Bob claims in terms of reasonable representation versus bounding, there's a wide range of variability of where our high uncertainties are. But the principal factors are the ones that are most important to performance, and are the ones that we're certainly going to spend our principal time on in terms of improvement. And that's what this testing program is laid out to do, you know, seepage, UZ flow and transport, drip shield performance, waste package performance.

WONG: How are you then addressing those factors which you have low confidence in?

YOUNGER: Well, I think maybe what you're getting at is the question of which ones will we try to bound with enough confidence that we can defend that bound, versus which ones will be treated with a reasonable representation. Is that--

WONG: Yes.

YOUNKER: I mean, on a case by case, I can't give you an answer to that, but I can say that that's that integration effort that's going on right now between performance assessment and the leads for each of the technical areas in trying to establish do we have enough information, is our uncertainty adequately characterized. But this is one where we will treat as a reasonable representation versus some of the other factors that will be treated as bounding, because we can defend the bounds, but we really don't have the time and money to put the full representation together, and we don't think we need to.

PARIZEK: Alberto?

SAGÜÉS: Let me tell you first that I appreciate all the time you have taken in fielding so many questions, and it's been a long presentation, so let me just say that I'm very glad to see that the program shares some of the concerns that some of us had about issues such as, for example, corrosion products that may develop over long time periods. Also, the attention being paid to natural analogs, and I sometime look forward to seeing the Titanium information that you're compiling. Of course, there have been compilations of the Titanium information, but especially I would like to see if you're developing some information on the performance of Titanium under varied conditions. That will be certainly something very, very interesting as it develops.

I wanted to call attention to one point in your transparency Number 24. That's something to put things in perspective, because I think that this brings up pretty much the kind of challenge that the program has to deal with, and those of us who review the program also have to deal with.

A statement is made there which is, you know, would appear to be a very reasonable statement. General corrosion rates are low, less than one micrometer per year. Now, for many applications, one micrometer per year or less is indeed a very low corrosion rate. But if we look at this in the perspective of the test, at one micrometer per year would mean one millimeter after one millennium, and it would mean ten millimeters after 10,000 years. And, of course, we're talking here about precisely that kind of time scale.

And then, of course, we only have two centimeters to deal with, and corrosion being what it is, the dispersion on corrosion is likely to be under the corrosion itself. So, you know, if the project were to demonstrate that corrosion rates are, say, one micrometer per year or less, that really would appear not to be enough by any means, because that means that the large fraction of the packages under those kinds of corrosion rates could very easily indeed be perforated after 10,000 years.

So I think that the meaning of the word "low" should be looked at in this context every time, and I'm sure that Joe Farmer is going to be able to address this. But we may have to talk like one-tenth of a micrometer, one-hundredth of a micrometer, or something on that order, to begin to feel comfortable about that being a low number.

FARMER: Just one comment, Alberto. When we look at the measured corrosion rates that come out of the long-term corrosion test facility, as you well know, the rates are so low that we're basically getting measurement error, and we can only bound what the upper limit is. It looks to us right now that somewhere between 95 and 96 per cent, looking at Alloy 22 as an example, 95 to 96 per cent of the measured corrosion rates based on weight loss appear to be below 150 nanometers per year, or .15 microns per year.

So we have actually four outlyer data points, and we're not sure if they're real or if they're just outlyers, and those four data points seem to be uniformly distributed between .15 microns per year and .75 microns per year. But certainly 95 to 96 per cent of those data points would indicate that you probably would have, you know, in excess of 100,000 years of waste package life limited by general corrosion.

And as, you know, you've also seen when you visited and were trying to use the atomic force microscope and other techniques to go in and make these measurements with much more precision and much better finesse than we've been able to do with the weight loss measurements.

SAGÜÉS: That's right, and that's a very good point. I wanted indeed to make sure that collectively, we have a feel for those numbers.

We also have in addition to the very long time, we have the very large number of packages, of course. So, you know, again if we say that maybe 5 per cent, in 5 per cent of the cases, the corrosion rates may approach or exceed that number, well, now again we have in these large numbers, fighting against us. And I just simply wanted to mention that I think that we all want to keep in mind the formidable kind of challenge.

PARIZEK: Bob Andrews. Do we have a few more minutes if we take a few more questions at this point? We don't have to meet with the public until 11:30. Okay. Well, we don't want to erode into your time schedule.

Okay, Parizek, Board. I have a few comments and questions, and one I share with Chairman Cohon. He indicated that the general presentation was well structured and shows a highly focused program, and we want to compliment the program for that. Your presentation reflects that, showing that you really have thought about a lot of these issues, and unlike maybe some people who come for the first time to these meetings, you get the feeling this might be a National Science Foundation random number of projects that need to be funded.

Rather than that, I mean all of the different things that are ongoing or need to be done have a purpose, and they fit into this grand scheme in a way that I think everybody should understand.

The question I have is whether the funding will continue in a way that allows us to progress in an orderly manner. Sometimes, it's a little hard to know what will be funded this year and what won't. For instance, I thought at Beatty we learned that maybe the Phase II Busted Butte experiments might terminate, and that either is a funding problem or maybe the relevance of those rocks to other rocks under the repository. So from time to time, we're not always sure exactly what will be funded and what won't be funded.

And part of this goes to Lake Barrett's presentation yesterday. You know, obviously if there's a cut in the budget, some things are going to have to be deferred, delayed, and again it's a little hard to make that judgment.

Site recommendation seemed to be a high priority, and with it is a lot of the efforts that you outlined for us. Can you make any kind of comment about that, as to what would drop out or have to be deferred?

YOUNKER: Yes, I can say that certainly at the planning level that we're at right now, which is kind of assuming that we'll get somewhere between the House and the Senate, I think that this work is solid and will be funded, the work that I've described. Now, of course, there's a question of how much of it, you know, how big is it, but the question of what happens if we come out toward the lower number, you know, I think Lake indicated yesterday, and maybe Steve as well, that

I guess we all know that that will be a different program. You know, certainly that number is low enough that we would have to go back and plan.

My personal view is because we would still presumably focus on what's important for site recommendation, these are still the tests and the analyses that will receive the highest priority. It will just be a question of how much are we still able to fund then at the lower level.

But I think unless it goes toward the lower number, I think this program that I've described is in our FY0 plans, and we expect to be able to cover it.

PARIZEK: Now, Chairman Cohon wrote a note to me saying what's the basis for anticipation that a realistic 3-dimensional flow model will be produced for the project? Again, that has to do with the saturated zone efforts.

YOUNKER: Well, and that one certainly is, you know, projecting a little bit further out in time to when we can get some results from an alluvial testing complex, you know, in cooperation with Nye County's work. So I think that one is just our hope that we have additional information, better hydrogeologic framework, you know, some additional geologic mapping that is being fed into the overall flow system modelling for saturated zone. Those are the basic reasons why we think that area is going to be improved.

PARIZEK: Now, as it relates to transport, that would be the Eh/pH work as well as the Kd work?

YOUNKER: Exactly. Yes.

PARIZEK: We understand a number of samples have been taken from the Nye County drilling project for sorption experiments in the lab. And I guess maybe there's a detail now that I don't know what's going on in that area. What samples are being included in those experiments? It's not clear to me what has been subjected to lab testing.

YOUNKER: I don't think we have anybody here, I mean, who will comment on that today, but that certainly is a topic that we could go into at another time.

PARIZEK: There's another concern I had with regard to the groundwater standard, you know, if we actually have to worry about our drinking water standard of the repository. Is there any effort being put into the possibility that might be required, and then what might come out of the repository other than radionuclides? Because it seems like all of the analyses aim at the radionuclide releases, but on the other hand, if in fact there may be another standard. Do we have any feeling of what other things should be looked at, or are being given consideration to make sure that you can comply with the drinking water standards?

YOUNKER: Certainly a lot of the background work that we've done as we've helped DOE prepare to comment on that rule has been looking at that, and I don't know, Bob, do you want to comment on that at all in terms of what other constraints it gives us if we have a drinking water standard?

ANDREWS: Well, actually I think EPA probably should answer that question, because I think what they brought into the 197 is only the radionuclide part of the groundwater protection.

YOUNKER: That is true.

ANDREWS: Not all other constituents like, you know, lead of chromium or whatever. But maybe they should answer that question.

YOUNKER: But in terms of what the drinking water standard dose is, though, I don't think that causes any fundamental change in the way we're going to model and test, you know, to do our performance analyses.

PARIZEK: All right, I have a couple more questions from Chairman Cohon, but I think perhaps we'll save them in the interest of time. Leon, did you have a question?

REITER: Leon Reiter, Staff. I want to venture into unknown territory called the waste form. And one of the most interesting things I saw in the comparison between TSP/VA and what the NRC had done had to do with dissolution of the waste form. It seems to me, if I remember correctly, and I stand corrected, they had a much lower rate of dissolution, and when I asked what was the reason for that, they assumed a different composition of J-13 water.

The second thing, they also presented possible models for which the dissolution rate could be even lower. Now, Bill Murphy presented a model by using Pena Blanca. I never heard this mentioned. Is this some sort of significant barrier that you're overlooking?

YOUNKER: I don't think so, and I have heard discussions about it, but I think I should defer to Bob. He can probably address that much more critically.

ANDREWS: Bob Andrews again. You know, in the VA, we did look at a number of alternative models for waste form degradation, one of which approximated, you might argue, what the NRC was doing with different groundwater compositions and reduction of rates in different groundwater compositions. That was not the base case in the VA. The base case in the VA was the more conservative, more bounded assessment.

We got the same comments from our own peer review panel, talking about the complexities associated with the chemical water/waste form interactions.

Right now, and I'm not going to speak to exactly what's going to be in the SR, but I think we will probably argue, and I can stand corrected a year from now, so don't take this too far, we'll still be using that bounded assessment. You know, the complexities and uncertainties associated with chemistry inside the package and its evolution with time, and that chemistry as it interacts with the waste form, and it changes with time, is just a very complex system with a lot of uncertainties in those models.

So it's in some ways going to be easier and more defensible to just bound it with the intrinsic dissolution rate, which is what the base case in the VA was. But we might change that, but right now, I would say that's probably what we're doing.

PARIZEK: I think we ought to go on with Bob Andrews presentation. Thank you very much, Jean, for a good discussion and a very clear presentation.

Bob will give us now a run-down on introduction to model validation, the processes involved. There are many models that have to be validated. We'll hear this afternoon two examples in more detail.

Bob is from the University of Illinois, as part of his training, and has a major responsibility for developing and documenting TSPA for site recommendation consideration reports. And everybody should know Bob, but he's already answered some of the questions that might come up, and some more of the ones that we had, we'll save for this afternoon that are kind of appropriate from Chairman Cohon and others.

ANDREWS: Your first question might be why is a PA guy giving a talk on model validation. You know, shouldn't it be some process level guy who's going to talk about the confidence in the model? And what we decided to do is kind of break it up into sort of introductory and why we care about validation, and sometimes I'll put it in quotes, and other times I won't, and then we'll follow this afternoon after lunch with two particular examples, one in the UZ and then one in the waste package, of the particulars of how in two particular areas, the process modelers are coming up with what they believe are valid representations of their particular components that feed into the performance assessment.

What I'm going to do in this briefing is to talk through a few definitions of validation just to put it on a common wavelength here, the requirements for validation. The word "validation" is not used anywhere in Part 63, the word "validation" is not used in Part 197. The word "validation" in fact was not used in Part 60 either. In some of the background documents to Part 60, the NRC had a lot of excellent dialogue about that particular word and how that word is used commonly in a scientific endeavor versus how that word is used in a decision making and a regulatory and a licensing kind of endeavor.

But the word "validation" still exists, and we want to talk to it and talk about what it means to us and what it means to the process modelers.

We'll briefly go through some general lessons learned from some international efforts, look at some perspectives that have come out, one is a very recent NRC combined White Paper, I think they call this, NRC, and the Swedish equivalent SKI, and the folks down at the center have a White Paper that came out in April on their definitions, if you will, of validation.

And then we'll talk about some general approaches to develop confidence, starting first with confidence in the safety case, then going to confidence in the performance assessment that supports that safety case, and then going down I think where the panel and the Board is most interested, and that's the confidence in the models that support the performance assessment that supports the safety case.

So if we go to the next slide, just a few definitions. First off, it's a comparison, you know, of the model, with some relevant observations, whether those are experimental observations which might be in the lab, or in the field, analog type studies, whatever the comparison is, is comparison of a model prediction of how a particular process is behaving, with direct observations related to that particular process.

This is coming from a quote from IAEA back in the early Nineties. A model is considered validated when sufficient testing has been performed to ensure an acceptable level of accuracy. Well, the definition of acceptable will vary, depending on the specific problem or the question being addressed or asked of that model. So the acceptability of the validity, if you will, is then tied to the intended use of that particular component, that particular model as used in some kind of application. The application of course we're talking about is those models as they're linked together to make some assessment of how we believe this system behaves or performs.

Also coming from another quote, which is somewhat subjective assessment, there's no objective determination that this model is valid. It's somewhat subjective based on the record, based on that the individual investigator, plus the reviewers of that individual investigation has come to, using all pieces of information to support that particular aspect of the system.

I do have in the back of the handout, the direct quotes from Part 63 and Part 197 on reasonable assurance and reasonable expectation, because that's really where validity or confidence comes in from a regulatory perspective, is in those two terms. And the direct quotes are in the back. These are just paraphrases that proof is not to be had in the ordinary sense of the word. EPA has required less than absolute proof, because absolute proof is impossible to attain.

You know, perhaps this is where our peer review was going, that in determining probable, where their definition of probable was an exact, precise prediction, it says that's impossible. You know, absolute proof is not to be had. There will still be retaining uncertainties, in particular over the time frames that we're dealing with. We just do not have direct observations over the time frame, or the spacial scales of interest.

And then they both acknowledge that there's greater uncertainties in making long-term projections. That's EPA's words, and NRC's words are demonstrating compliance involves use of complex models that are supported by limited data. You can't exhaustively test every single component of every single model that's used in the performance assessment.

DOE brings forward some of those concepts more from a quality assurance perspective is where model validation comes in. Here I'm quoting from the most recent version of the QA requirements document, DOE document.

Models shall be validated to a level determined by the intended uses. Well, that's really why I'm up here, because the intended uses of the models that Bo is going to talk about this afternoon on UZ flow and that Joe is going to talk about on waste package degradation, the intended use is to make an assessment, to make prediction, if you will, with uncertainty of how we think this system performs.

The intended use of that UZ flow model is not to exactly evaluate the exact quantity at ever square centimeter of rock or within every fracture within the rock. The purpose of that UZ flow model is to evaluate globally the average percolation fluxes through the mountain, and on average, how that percolation flux is distributed between the fractures and the matrix, globally how seepage behaves, not exactly where you might expect to find seeps within the nearest square meter or for ten square meters.

So the intended use is more of an average approximation. It's not the exactness of a particular flow path or a particular velocity that that model is being run.

And the same is true of the waste package degradation model. The intended use is not to say exactly which package failed and exactly how that package failed, but within the 10,000, roughly, packages that exist, what's the likelihood of some packages failing. When they do fail, what's the general morphology of that failure in terms of the total surface area exposed underneath that opening.

So intended use of the models I think always has to be kept in mind. The intended use also incorporates that those models will be used in a probabilistic sense. The uncertainty in those models, the uncertainty in the parameters in those models will be captured to the best of our ability, or bounded to the best of our ability. And that's the intended use.

So taking Leon's example, you know, from earlier on waste form, which is not one of the ones of subject discussion later on this afternoon, the intended use is just to find how many nuclides came out into, in this case, a liquid phase, as a function of time, given the environmental conditions that exist in that package. It's not a precise number.

There is a huge amount of uncertainty and complexity, probably 20 pages of that complexity mentioned in our own peer review report on waste form, water, chemistry interactions, and the lack of detailed information on that. So it's just much easier to go in there and say that one I'm going to bound. I'm going to defend that bound, et cetera.

The QARD also acknowledges that the validation will be accomplished by comparing the analysis results against data acquired from lab, field, natural analogue or subsequent relevant observations. If you don't have any data from any of those sources, it says use an alternative approach. One of the alternative approaches is a peer review of that model, that component of the assessment. But generally, and I can't think of any area where we don't have some technical information, some data, whether it be laboratory data or in situ data, and in many cases, analogs that support the models that are being used.

Okay, going on, the international community has worked on model validation for the last decade and a half, or so. In fact, it started before the time frames I have there, but the earlier times were more focused on software, focused on code, comparison, comparison of different codes. They quickly realized that it wasn't codes that were the issue. Generally the codes, if one had the same conceptual model and was modelling the same processes, the codes were more or less given the same answer. You know, you could have pulled off the shelf petroleum reservoir engineering code from Company X, and flow and transport code from Lab Y, and gotten the same result. And that did happen, you know, lots of times in the mid Eighties.

The issue was in the analysts. The issue was in the data and the conceptual understanding as one applied that piece of software. So essentially, there's about four, and there's probably some that I'm missing here, and I apologize to any who might have been involved in others. One related to flow and transport type models, one related to geochemical models, one related to vitrosphere models, and one related kind of to near-field models.

To the best of my knowledge, there's no international model comparison of waste package materials, waste form type models. So you're hitting the natural system type models and the biologic system type models.

But these have been going on for a number of years.

I tried to summarize the lessons learned very simply on the next page. It's kind of difficult with the wide range of studies, wide range of principal investigators, a wide range of countries and analysts. Each of those validation studies looked at, you know, ranging from five to tens of example field type locations or test locations where, you know, five or ten groups would look at their models and try to explain the observations using their models. So making their assumptions, incorporating what they felt were the right processes in their models, and then trying to assess by comparison to direct observation whether that's the field. Many times in situ tests were used as the comparison basis.

What do they conclude? Well, validation is difficult. So in many cases, different analysts, different groups, looking at the same test configuration, trying to interpret that test and compare the results against the results of that test, they came up with slightly different results. So it's a difficult task.

Why is it difficult? Well, in some cases, and this is their kind of assessment of their own validation efforts, and I think there's some people on the panel this afternoon who were intimately involved with some of these. I know Chin-Fu was and I think others were, too. So they can probably talk to their own experiences associated with these international validation efforts. I don't know if there's any NRC people on the panel this afternoon who were directly involved with this, too. So they can give you their own read, and it might differ with these, and that's cool.

But there's a thorough understanding of the processes. If you didn't factor in a process into your model, and that process was in fact driving that test, then clearly you had some difficulty in explaining the results of that particular test. That was especially true in a number of the flow and transport studies done earlier, some of the work, there were actually processes in and around the drift that the models did not have in them, some of the coupled processes that the models didn't have in them, so they didn't explain some of the observations very well.

They did acknowledge that some comparison with experimental results, and this might be laboratory results, did enhance the confidence in the models. In many cases, detailed comparison with the tests, detailed comparison with point values from the tests, was very difficult to achieve. But some integrated--and I used the word performance measure here, that might not be very precise--but a little more integrated measure of that test was reasonable to achieve.

You know, it was difficult to achieve exactly where water might be dripping, but reasonably, most people were able to predict how much water was dripping. So there's a distinction between, you know, the precision or location or accuracy versus some average characteristics of the system.

And they acknowledge that by comparing different conceptual models, even the same analysts comparing different conceptual models, it gave useful insights into the validity of the models for their intended purposes.

Switching gears from the international to the recent NRS/SKI White Paper, just a few bullets to try to capture the main essence of that White Paper. First off, a point we've made already is the level of confidence required for model validation or for a particular model is tied to the importance of that model in the decision making process. You know, if the model is less significant, less important than the degree of validity or the degree of confidence, you know, one requires in that model is somewhat less than something that's of major significance to the performance or to the decision making process.

They also go on to say, not surprisingly, considering the words I gave you earlier about reasonable assurance, that exact prediction is neither expected nor required. Goal is to establish the adequacy of the scientific basis and demonstrate it is sufficiently accurate for its intended purpose.

They go on with, in the next slide, with an example, I think they call it a validation strategy of the steps that in particular NRC and SKI would expect to see in a normal application of developing confidence of the application of the models, starting first with a compliance demonstrate strategy, determining the goals, determining the existing degree of validation, comparing the goals with the existing degree, deciding whether to revise the strategy, and then finally obtaining additional information.

If I go to the next slide, I make an attempt to compare those steps in the strategy with what I would argue is DOE's implementation of that strategy as we laid out yesterday for you, and as was laid out in fact in the VA for you prior to the NRC/SKI White Paper being released. And quite frankly, as I was looking at this last night one more time, I realized I probably should have broken this DOE implementation up into the VA versus the SR, like I did yesterday, because there's different references I would have used for the VA implementation of effectively this strategy from the SR implementation of this strategy. So I'll walk through that as we go.

First, define the compliance demonstration strategy. Well, that's what both Abe and Mike Voegele presented to you yesterday. The compliance demonstration strategy is, in DOE's parlance, the repository safety strategy. The repository safety strategy is in Rev. 3 in draft form now, looking forward to the SR.

In the VA time frame, it really was captured in

Volume 4 of the VA. There was a repository safety strategy that went hand in hand with Volume 4 of the VA, but they were consistent and had the same information within them.

The goals for model validation, i.e. how much validity--by the way, you won't find the word "model validation" I don't think in VA Volume 4, nor will you find the word "model validation" in the repository safety strategy. But in both cases, they talk about confidence in models, or uncertainty in models. So confidence is like validity, and uncertainty is like one over validity.

So you'll find the same, or one minus validity, I'm not sure, you'll find the same thought process in Volume 4 of the VA and in the repository safety strategy without using the terminology.

So the goals for model validation, there's tables in Volume 4 of the VA, and the repository safety strategy, in the very fact that it's somewhat divided between principal factors and factors, is really defining the goals with respect to the significance. And that significance has buried in it already the uncertainty in that particular factor. So it's somehow embedded qualitatively in that factor. And of course in the ultimate SR and VA, it's in there quantitatively. But in the repository safety strategy right now, it's in there qualitatively.

Determine existing degree of validation. You know, the Volume 4 of the VA gave, in those tables, gave a somewhat qualitative, subjective, because remember validation is subjective, assessment of the degree of validity of each of the component parts used in the TSPA/VA. Some things we had a higher degree of confidence on. Some things we had a lower degree of confidence on. I think that high degree of confidence/low degree of confidence was more or less endorsed by the peer review. They might have differed in a few areas, but we said, you know, cladding was probably of moderate to low confidence, and I think the peer review probably said low to very low. But it was close to the same order of magnitude.

The next step is to compare the goals with the existing degree of validation. Well, the Volume 4 of the VA did exactly that. It said here's my goal for the degree of validity I think, or we, the DOE, thinks is needed for that component of the system, based in part on its significance to post-closure performance, and here's my current confidence level and, therefore, here's what I think I need to do. So that comparison really was in tables within Volume 4 of the VA.

The decision point then comes after the VA and the project officer went through that decision point of whether to revise the compliance demonstration strategy. One part of that revision can be go out and get additional information to remove some of that uncertainty. One part can be go revise the design to accommodate some of that uncertainty. And, in fact, the project did both of those avenues. It did revise the design, and it did update or is in the process of updating the strategy to reflect that new design.

Oh, here's the other one. Obtain additional information to support the validation. So for those things that are still important, for those things that still need to be of sufficient confidence for the intended use in post-closure performance, go out and gain additional information. And I think Mark Peters yesterday afternoon, and Jean this morning talked to those areas where the project is focusing its resources to do that additional information with respect to this strategy.

So in a way, you know, this strategy, the validation strategy, as implemented, is implemented within the repository safety strategy and all the supporting analyses and documents that are behind the safety strategy.

Okay, other people have had some insights with respect to model validation. The TRB tried to capture here a few of those--I'm not sure whether in TRB reports the word "validation" explicitly is used, but I'm sure the word "confidence building" is used frequently throughout the reports.

It's acknowledged in some of the TRB writings that to make robust decisions, and at each step, decisions are being made, there's decisions made on the sufficiency of data, sufficiency of models, sufficiency of analyses, including PA analyses, sufficiency of the safety case, and ultimately, you know, the sufficiency of decision, sufficiency of the information to support a decision. And that's not only technical information. There's a lot of other inputs into that decision, clearly, as the Board has pointed out numerous times.

But the technical side acknowledged that first, these robust decision can be made if the uncertainties are fully and accurately addressed, so we acknowledge them, address them, evaluate their significance to the performance assessment, to the safety of this system. Carry out those sensitivity studies using different assumptions, and show compliance with a high degree of margin. So those three aspects would allow one to make more robust technical decisions.

Identify how the PA conclusions will be used to make those decisions. And I think we talked about that a little bit yesterday with respect to the sensitivity analyses, the uncertainty analyses, et cetera. And make sure that the PA is as transparent, I would add as possible--maybe you wouldn't add that word--you'd just say make it transparent. Make sure the assumptions, their basis and effects are clearly and explicitly stated, and you'll get to that this afternoon with two of them on UZ flow and on waste package. Make sure the key parameters are traceable and make sure that TSPA has undergone an independent review, which of course the VA did undergo.

Now I'd like to shift and talk to kind of from the top down, and as an introduction more or less to Bo Bodvarsson and Joe Farmer this afternoon. And the top down is having confidence at each stage of the decision making process, starting with the safety case, going down to the performance assessment that's a part of that safety case. It's not the only thing in that safety case, but it's a part of it. Down to the models used in the performance assessment, and finally, down to the data and information used within the models.

I'm just going to give some general words here. Bo and Joe will talk this afternoon essentially about this one, and with probing, I'm sure you'll get down to this one that supports this one, confidence in the data and information to support their models.

Starting with the top and going down, the general approach to developing confidence in the safety case is what Mike and Abe talked to you yesterday about. I mean, the repository safety strategy lays out DOE's approach to having confidence in the overall safety case, but it's tied first to the robustness of the system, which you could say are directly related to the TRB insights that we had on one of the previous pages, and it's also tied to the quality of the assessments used to support that robust system.

So it includes a well defined PA approach, component models that contribute with a high degree of confidence, relevant data have been considered, and result are fully disclosed and subject to QA and review.

So these words are in part from the repository safety strategy and they're in part from the OECD/NEA White Paper on building confidence in safety assessment. But they're the same words.

The next step below the safety case is the actual performance assessment conducted in support of that safety case. And there, kind of the steps or the approach is to first identify the levels of importance of the individual components that affect long-term safety, identify the degree of validity in those component models. This really goes down now to the next level below, because the confidence in the models is down at the process level, the confidence in how those models interrelate is at the TSPA level, and how the inputs from one go into the--or the output from one go into the inputs of another.

Identify the full suite of reasonable alternatives. You might classify those as features, events and processes that are either included in the analyses or explicitly excluded from the analyses, and the basis for their exclusion is documented and justified.

There's screening of the features, evens and processes, and there's also screening of the individual sub-component or sub-system or component models to determine those components of a model that need to be carried forward into the assessment of performance.

The next page, not only are there models in the application of the performance assessment, but there's parameter values within those models. There's as much, or needs to be as much scrutiny on the parameters within the models that are used and abstracted and incorporated in the PA as there is in the models themselves. So there can be sub-system or component screening of parameter uncertainty, and the significance of that parameter uncertainty, and which parts of the parameter uncertainty need to be directly incorporated in the performance assessment.

Finally, there's an evaluation of the system performance to the effects of those uncertainties, and this in part is to help evaluate quantitatively the barrier importance of individual components of the overall system.

And, finally, last but definitely not least, is to document all of the above in a manner that allows one to transparently and traceably see how the conclusions were reached.

The next page was in there for the graphical picture of developing confidence from the data up through the TSPA. It's from yesterday. We can skip over that relatively quickly and go on to more or less the last introduction to this afternoon's talks, which is the approach to developing confidence in the actual models that are used within this prediction of performance.

We talked yesterday about a wide range of models. There's something like 40 analysis model reports that are directly fed into TSPA. Mike Lugo talked to you about a total of 168, I believe, analyses and models that support those. So it's those that we're talking about, and I think Bo has probably, correct me if I'm wrong, 30 of them, and Joe Farmer has 20 of them. So you'll be talking to those 50 this afternoon, or a subset of them, depending on how much time we have.

But in general, the confidence building in the models themselves is based on their comparison to direct observation, laboratory observations, field observations, analog studies as appropriate, and some peer review if appropriate, if there's no other source of information.

And I want to say the appropriateness of each one of these sort of depends on the type of model. You know, for Bo, he'll talk more about field tests and a little bit about analogs. For Joe, he'll talk more about laboratory experiments. So the type of information used to support the validity of the model really does depend on the model.

In conclusion, all I'm up here for is to kind of introduce this afternoon. But validation is a process, you know, for providing increasing levels of confidence as one goes through a decision making process. One gains information. It is the scientific method, if you will. One gains information, one tests that information using models. One revises models with new information, et cetera. But it's a process that one goes through. There's no black and white, yes and no. There's varying levels of confidence. Those models as they're incorporated, incorporate that uncertainty as appropriate.

The second point is that the model validation approach that the NRS and SKI laid out in their White Paper really is more or less what the DOE is following. DOE calls it something slightly different, but it is more or less following those same six steps in the approach laid out in the White Paper.

And, finally, as I've said several times, Bo and Joe will talk in much more detail about their particular parts this afternoon.

So with that introduction, Dick, I'll turn it back to you.

PARIZEK: Thank you. Questions from the Board? Chairman Cohon?

COHON: Thank you. Let me offer, suppose you had two different goals for your model, for a model. One is to estimate the expected value of dose, and the other is to estimate expected value of dose and the variance of that dose. Would you expect that that would have different implications for validity of the model and underlying models?

ANDREWS: Well, first off, as soon as I determine the expected value, I'm going to have the variance around that anyway, because the expected already is a mean, and has a variance around that.

COHON: Right.

ANDREWS: So I can't--

COHON: But what I meant by this, and I should have been clearer. Suppose the variance of the dose was a decision criterion as well as the expected value of the dose. Do you think that would have implications

ANDREWS: I think so, yeah. I think I would--I'd have to think through how those models are incorporated, and we are incorporating the uncertainty in those models to get that expected value regardless.

COHON: So the question is whether you would do it differently if the variance was also a decision criteria.

ANDREWS: I don't think--

COHON: Or would there be a higher level?

ANDREWS: I don't think dramatically differently. I mean, we'll be coming up with an, if you will, a PDF on dose, you know, over the 10,000 year time period. There is a point on that PDF called the expected value. But the full PDF will be there. It will be there as part of the analyses. I think that it's the same, and whether the regulation, you know, the old 191, asked for a CCDF of releases, you know, at the accessible environment boundary, that had to incorporate uncertainty in the models and uncertainty in the parameters into it. And what we're doing is not dramatically dissimilar from that.

COHON: Okay.

ANDREWS: I don't think it changes really, and now you're going to throw me the next question and I'm set up here.

COHON: No, this is an honest question. I tend to give you a hard time only because I find your presentations so clear and they prompt, they stimulate questions in me. And your answers are always very good. This is not patronizing, and I'm not setting you up. I promise.

Suppose your decision criteria were expected value variance and the confidence, quantification of confidence in your estimated of expected value in variance, so you're have three or maybe four criteria. Do you think that would have implications for model validity?

ANDREWS: Yes, there I would, because I think there will be areas where we will go in with what we believe is a demonstrable and conservative bound, and we won't test every bound, and it's the range within that bound, and it's significant, which if you wanted that last step, the confidence level, I think you would want to do that. You'd want to really incorporate every part, and the full range of every part.

COHON: Thank you. On Slide 11, you talk about the very first sub-bullet under more robust decisions, uncertainties are fully and accurately addressed, and of course we all agree with that. I would like to see, say, fully and accurately addressed and communicated.

There's an issue here of whose decisions we're talking about. I'm confident that the program will be addressing these uncertainties to support the program's decision making, but I think that your understanding of those uncertainties also have to be communicated to others who have decisions to make, including this Board and political decision makers. That wasn't a question.

Finally, just sort of a semantic discussion, which I think is more than semantics, I have a problem with the idea of degree of--the degree to which something is valid. To me, validity is like perfection, either valid or not, you're perfect or not. But we all know it's incorrect English to say more perfect, less perfect. Degree to which you are perfect, the degree to which you are valid.

Now, the reason I think it's more than semantics, though, is that it seems to me that I liked your structure very much. You have to understand the goal for the model, the role that it's playing, and what we demand of the model, and on that basis, and only on that basis, can you declare something valid or not? The degree to which it's valid, to use your phrase, really is a statement of our confidence in its validity.

So it seems to me that what we're really after is a statement that it's valid for this purpose, and my confidence in that claim is this. Am I off base here, or is that consistent with what you mean by degree of validity?

ANDREWS: I think the degree of confidence, can you have a degree of confidence? And I equate confidence and validity as synonyms, and if I can have a range of degrees of confidence, then I can have a range of degrees of validity.

COHON: So this is what you really mean by degree of validity.

ANDREWS: Yes.

COHON: It is the model is valid for this purpose at this degree of confidence.

ANDREWS: Right.

COHON: Okay, thanks.

PARIZEK: Paul Craig?

CRAIG: Craig, Board. This is in a sense a follow on to Jerry's comments on variance and margin of safety. As you were talking, I was thinking that I hope I get to fly home at some point. Maybe I will, given the storm. And I hope the plane will work right.

There are a lot of subjective elements that go into this, and your presentation made that very, very clear. How good is good enough, is what we're talking about. And what I'm concerned about here is the level of confidence the user has in the whole process, some ultimate user, in my case, the person who's going to fly on the airplane and hopes to get there, and what I'm concerned about is the difference between whether something will probably work versus the idea that it will work with a really high level of reliability. If I thought that the airplane was only going to probably work, I might decide to take the train.

Now, when we look at the regulatory perspectives, which you have here, they don't seem to be very concerned about a high probability of it working. They use these words "reasonable assurance" and "reasonable expectation," and you properly labelled those a discussion on acceptable level of accuracy.

And so what I'd like to get us to do is to reflect a little bit in the context of our expectations for this 10,000 year or more performance of Yucca Mountain, whether reasonable assurance and reasonable expectation is really what we're after, or are we after something substantially more than that? And perhaps that's what the Board is getting at when it talks about, as shown in the slide that was up there just a moment ago, as going beyond the standards in order to enhance confidence, or going one step beyond, meeting the standards robustly.

But what I'm really focusing on is the difference between reasonable and high confidence, if there is such a difference.

ANDREWS: I don't know if there's a difference of not, Paul, quite frankly. Maybe I should stop at that because I can see my mouth opening and inserting a foot. Maybe somebody from a more regulatory background than I can talk about reasonable assurance and reasonable expectation versus--I mean, I think varying here is scientific--I have the full quotes at the back. You know, there's a scientific, they don't use the word validity, but scientific confidence in the underlying assumptions, underlying assessments, the underlying judgments that had to be made by the analysts as they applied limited information, and it will always be limited information, limited base, limited time, as they apply that information to their models for the intended purpose.

You know, Abe, if you want to add something to get me out of this jam here?

VAN LUIK: Yeah, I was not going to shed light on this, except to ask for a clarifying statement. When you get on an airplane, don't you have a reasonable expectation of getting home? Otherwise, you wouldn't have gotten on the airplane. And I think it's an individual interpretation of what those words mean. If I wasn't reasonably sure that this airplane was going to take me home, I wouldn't step foot in it, and I think if we are--and the key is reasonable. If you have an unreasonable fear of flying, none of this applies. If you have an unreasonable fear of DOE, you will never have confidence in anything that they do.

So I think, you know, what we're talking about here is your individual interpretation of what is reasonable or unreasonable.

HANAUER: My background is in nuclear power plant safety, and reasonable assurance is intended to be a very high standard, in spite of what the dictionary might say about the word reasonable, and in spite of what Mr. Clark said yesterday. I sign a lot of ACRS reports to the Chairman of the Atomic Energy Commission, as it then was, and the conclusion was that we found reasonable assurance that the proposed plant, or the operation of the plant as built, would not cause undue risks to the health and safety of the public. And we intended that to be a very high degree of assurance.

PARIZEK: Priscilla Nelson?

NELSON: Hi. I'm recently having a lot of conversations about model based simulation of performance, and as an interactive, what you might call some aspect of validation, is a two-way street where a model feeds back into the experimental environment, which feeds back into the model, increasing the confidence in the model. And it seemed like this discussion was very much one way, with the experiments putting into the model rather than having the model feed back into the experimental scenario. So that was one observation.

I think another observation that I had just from my perspective would be I'm not sure what I'd do with, for example, if you had two models that we're try to, like for example equivalent and continuum and fracture flow, where it may well be that the input data are so different in character, and what you know about that input data is so different in terms of quality perhaps, or confidence, that it becomes very difficult to talk about, you know, validation of one or the other, and what you do about the two.

It's sort of the second observation that I'm not clear about after your presentation. And the third one is about the prospect of if you validated the models, such as Joe Farmer and Bo are going to talk about this afternoon, is the compounded model that includes those also validated? Or how do you investigate that?

ANDREWS: Okay, I realize those were observations, Priscilla, but let me assure you that trying to combine the first two observations, although I might have looked at this linearly, you know, do a test, do a model. In fact, it is in reality a very iterative step. In most tests, before the test, there's a model. In many cases, not all, in many cases, that model is a quantitative model, you know, assessing pre-test what you think you're going to observe, and the timing and frequency that you need to observe the things that you're going to observe.

That model then, once the test is ongoing, is compared against the actual observations, and in some cases, modified. That might be called a calibration step, you know, of the model rather than the model being applied in a direct predictive sense. But then the model is applied to predict the next phase of the test. So it's iterative between model test, model test, model test.

NELSON: That's more of an update sense, rather than have the model feed right back into the experimental environment in terms of defining what the experiments ought to be, and what the data acquisition ought to be. It's much more of a two-way thing.

ANDREWS: Well, I think in reality, it is a two-way thing.

NELSON: Okay.

ANDREWS: And if I take the example, and maybe Mark Peters can chime in here, but if I take the example of the drift scale test, large scale heater test, there were a number of pre-test predictions of that test. There are a number of predictions going on during the test. There is a decision to be made that those models will help make. That decision to be made is when to turn it off and when to lower down the power output, or increase the power output.

That decision--I think it's going to be lower, not increase--but that decision point will be in part based on the models, and the models saying this is a reasonable time to stop that test, because I've maximized the utility and the spacial extent of that test for the purposes of that model.

So the model is used beginning, in the middle, and at the end, you know, for real decisions on real tests. The same thing is probably true, although I can't speak to it as well, is the cross-drift testing. I know, or am pretty sure the LBL folks have done a lot of pre-test, and LANL has done pre-test predictions of what they think they're going to observe. And in fact those pre-test predictions will help to design the actual test layout.

So, you know, I think it does happen. Maybe we need to portray it in that sense, you know, as a confidence building conceptual pre-test, test comparison back of test against the pre-test to show people, you know, that there's continual learning and updating and revision, modification of the actual models.

NELSON: It seems like this will get you closer to have a site specific tool, where, you know, it's the general concept of a model is, to me, you're going to validate it for the experiment specific and the site specific data input and processes that you modelled.

ANDREWS: Yes.

NELSON: I mean, it's a very focused validation.

ANDREWS: Yes, it's focused on that hunk of real estate to which those stresses have been applied. And that's what you can do. You cannot stress the whole mountain. You can stress this hundred cubic meters of rock. And that's what you do and compare it to the model.

Your third observation, if I can jump to that one, the actual intended uses over spacial and temporal scales, the exact test does not capture. Clearly, we're looking at 10,000 years, and we're looking at spacial scales on the order of hundreds or thousands of meters, not meters to tens of meters. So there's always a--and that's I think the point in one of those, you know, validation lesson learned, was some integration of performance, if you will, provides a little higher degree of confidence for the model for its intended use than a direct comparison to specific test information.

But the hooking up of the models, you know, that I talked to a little bit yesterday with kind of a sub-system performance evaluation that you could compare those right back to, you know, the model output. You could compare those things.

NELSON: But I could imagine some cases where they're not independent models, where there is model interaction.

ANDREWS: There's a lot of model interaction.

NELSON: A lot of model interactions. And, therefore, the exercise of validating a combined model is different from one of doing one of the individuals.

ANDREWS: That's true.

NELSON: How do you do that?

ANDREWS: You turn off some of those interactions and make sure that at least that part of it works. You can only look at how information flow, how mass flows and water flows and nuclides flow through the system in making sure you are conserving mass and water and nuclides. That you can do.

NELSON: Thank you.

PARIZEK: Debra Knopman?

ANDREWS: I think Joe wants to add something.

FARMER: I'd like to make one comment about integrated models, because that's a situation we have with the waste package. And I think in our particular case, we measure thresholds, which Bob's group uses these thresholds as switches to switch from one failure mode to another. So we actually do have specific testing where we go in and make sure that these switches are appropriate, and that the thresholds for switching these modes of failure on and off are correct.

So I think there are some ways that we can go in and test and validate these integrated conceptual models, if you will, and we're trying to do that.

KNOPMAN: Knopman, Board. Insofar as your, I think, the program is trying to focus on site recommendation, and the decision making environment that you're going to be operating in there, are you or is it being contemplated, or have you already or are you contemplating doing some elicitation or interviewing or some discussion or focus groups with your decision makers, both at the departmental level and in Congress? Because I'm not so sure there's folks with technical training, and legions of papers have been written on the subject of differences of risk perception between technical audiences and lay audiences, and I'm not sure you--I haven't heard it yet in any of the presentations that there is an appreciation for how this question of how good is good enough is in fact going to be processed and dealt with in the decision making arena you're actually functioning in.

I think you'd learn a lot about it, and I think it would influence the research agenda, and certainly the way you piece together your safety case.

ANDREWS: I agree. I don't know if DOE, Abe or anybody, wants to comment or respond.

VAN LUIK: Abe van Luik, DOE. That is an excellent point and it's an excellent suggestion. What we have done is we have paid attention in a lot of meetings with different people with different viewpoints, and in fact, you know, some of the things that we know are not very important to performance, we intend to keep monitoring them, because they are so important to people's perception.

On the other hand, we are trying to make an effort to focus and close a program to answer a question and move on, so there's attention between those two, and your idea of perhaps investigating this with some focus groups is an excellent idea. Frankly, I hadn't really thought about doing that.

PARIZEK: Dan Metlay?

METLAY: Dan Metlay, Board Staff. You have made the point several times that the level of validity/confidence in a model is related to the decision to which that model will be used.

One could argue that the site suitability decision is in some sense less consequential than the NRC licensing decision, and therefore, one needs less confidence and perhaps by extension, less validity in the model at site recommendation than at licensing.

But the converse argument could also be made, that the most consequential decision is the site suitability decision and, therefore, more confidence is needed at that point than perhaps at any other point.

I guess I have a two part question. First, to what extent are different levels of confidence going to be attached to site recommendation and licensing? And since we've talked about confidence in a metric, how much difference will there likely to be?

ANDREWS: I guess I'm the point guy on this question. But I'm going to turn it over to Abe probably in just a second.

Our perspective is, you know, both decisions are very crucial, hard, scientific, technical, sociopolitical decisions. A lot of inputs into both of those decisions, I've talked to just one technical aspect of the decision with respect to scientific confidence in the analyses and the models, and the full suite of analyses and models going actually down to, you know, their scientific basis will be discussed in more detail this afternoon.

So both decisions have that same degree of scrutiny, of test, if you will. I think there are--now I'm going to speak a little bit for myself, so somebody from DOE probably should talk up. The amount of data, Mike Lugo went through yesterday the qualification aspect, you know, the data qualification from an NQA1 regulatory perspective at the different phases of the assessment, you know, 40 per cent at Rev. 0, 80 per cent at Rev. 1, 100 per cent at LA.

As one goes through that process of making sure the data are qualified from an NQA1 perspective, and the models are qualified and the software qualified, some additional bounding may occur between the SR and the LA based on the SR analyses and based on the safety case that's written after the SR analyses are completed. That's not to say it's any more defensible.

It's just that probably some of the data sets that may be difficult to qualify, you might want to remove that as an issue of concern to the regulator between the SR and the LA, and go in with even more bounded analyses for certain parts in the LA. That's a decision that's TBD. You know, I don't want to say that's a firm decision, and maybe Steve or Abe would want to tackle that same question. Or maybe we'd like to break.

PARIZEK: No, we can't take a break.

VAN LUIK: Abe van Luik, DOE. I think Dan brings up an excellent point, in that the audiences for these two decisions are very different. And, in fact, I think we are much more comfortable with a very technical audience such as the NRC presents than we are with the political decision making process which will be the SR's challenge. And I think when you look at that, the degree of confidence that we need for both is probably comparable, but the way that we present it would be different.

We can talk very technical and very detailed to the NRC, but I doubt if we can convince a congressman with, you know, how high the footage is on the documentation that we bring in. With a congressman, we have to make arguments that sound plausible and reasonable.

And so I think it's the way that the confidence is presented that's very different, but the degrees of confidence are probably comparable. And the original degree of confidence that we had when the two documents were very close together would have been exactly the same. But it's a difficult issue. It's the packaging for the two different audiences is different.

METLAY: Can I just follow up with a real quick followup question? You cited some what you called insights from the NWTRB on one of your slides, and one of the comments that the Board had made was noticeably absent in that, and that was the notion of establishing beforehand sort of standard of confidence. And sort of the analogy I've used in the past is shooting an arrow at a barn, and then placing the target around it and declaring I've hit a bull's eye. And it's a lot easier to understand confidence if one knows what the target the DOE is shooting for ahead of time, rather than possibly after the fact.

And I'm wondering what the DOE's thoughts are with respect to confidence, both in terms of some of the parameters that Chairman Cohon mentioned, the expected value of the variance or the level of confidence. Will we hear about that ahead of time, or just after the fact?

VAN LUIK: Abe van Luik, DOE. This was Bob's viewgraph. Why am I answering this question?

I think the reason that we left--we were very well aware that that was the TRB's suggestion, comment, and a serious one. I think the reason we left it off is because we're talking here about validation of models.

One of the internal requirements for applying the QA definitions of validation is to define a goal, state how close we are, exactly the same as with the NRC and SKI, define a goal, state what our current position is, and what we're going to do to get to that goal.

So at a technical level for a model, yes, we will do that. The overall statement of confidence on our total system performance assessment is something that we will stipulate what our confidence is in the TSPA/SR and the TSPA/LA. But as far as saying up front what that is going to be, I wouldn't even know what language to conjure up to explain what that would be.

So at a lower level, yes, we plan to do that. At the top level, we have to basically meet the legal regulatory requirements with sufficient margin that we feel comfort in the case that we're making. We are not going to get on this airplane without ourselves having a reasonable expectation that it provides public safety.

PARIZEK: Parizek, Board. Just one brief observation about this idea of prevalent expert judgment. When I don't have any data, I don't have any models, I don't understand the process, and I bring in expert judgment, and there's a risk to that, because that leads to the idea, like at West Valley, the distance of travel ground water will be 2,000 feet, when in fact that probably means it's only six feet. It's not permeable at that time with the ability to measure it, or there's no water table because we can't define it. We don't know how to define it. So there's always these things in the audit after that come back and says, well, it's the best we could do at the time, that's all we knew at the time, seems to be always a risk when you go to experts.

It's much harder to compare experts' opinion than it is maybe models. You said we could take the same codes, different people can produce a similar result. We can compare codes that come out kind of close by, and feel pretty good about that. But experts flaunt around a little bit. If they're noisy, maybe they're good. If they're not so noisy, maybe they're better.

But this probability distribution thing that we deal with, how is the program going to deal with the expert judgment? I know there's a whole protocol for doing it to make it reasonable. And maybe, say, you have to go on with the program and make hard calls when you have to make them, but it seems to me it's even harder to deal with that one than it is maybe some of the models and codes that we have to look at.

ANDREWS: Let me try something. Those aren't my words; those are NRC's words. But I'm going to get a distinction between expert elicitation, the formal process of eliciting experts that may in fact synthesize lots of pieces of information, from lots of different geographic areas and lots of different process understanding, to a particular problem with somewhat limited information.

You know, an excellent example and, you know, how we're still using them is in the seismic hazards and volcanic hazard assessment, using site specific information in both cases, but they're extrapolating that significantly, you know, to make an assessment of probability of occurrence.

I think what this is getting at, quite frankly, is the judgments that really do occur down at the analyst level as that individual is doing their analyses or developing the details of their model. There's judgment involved in the gridding, you know, of a UZ flow model, tremendous judgment of how to scale properties to the scale of the model when you don't have direct observations at the scale of the model.

So I think what this is getting at is the judgments that the analyst or modeler is making, you know, have to be acknowledged. I think we have excellent analysts and excellent modelers, and Bo and Joe will talk about some of them, who are using professional expert judgment in some of the details of their analysis. That judgment, of course, the review is checked, it's reviewed, it's synthesized in the PMRs, but it still will remain in any of these things.

So I think I made a distinction between elicitation process and what really still will be a large amount of expert judgment by detailed experts who will be on the stand some day to defend their judgments.

PARIZEK: Thank you. We have to go on with the public comment period, and we've taken some of their time. Thanks again, Bob.

COHON: Thank you, Richard, and our thanks to Bob and Jean for a very good morning so far.

We turn now to the public comment period. Let me first call on Walter Matyskiela. I probably still butchered your name. At least I attempted it this time. You might state it again for the record.

MATYSKIELA: This is Walter Matyskiela. People have been encouraging me to talk, so I'm going to make a few comments. I also would like to compliment the speakers this morning. I think they made very crystal clear arguments regarding the plans of the program and the issues.

But I think several people began to raise what to me is the more fundamental question than validating codes or models, and that is the idea of concept validation. To me, this program illustrates sort of a fundamental failure of the systems engineering process, as most people believe it ought to be practiced in the world, wherein you're supposed to identify the primary factors affecting the issue at hand.

In this case, the program has steadfastly ignored the issue of the heat affecting the rock, to the extent that we now have some examples that I'd like to give you that are reasonably absurd. We have, for example, a bunch of tests that have been done at Busted Butte on rock that is only remotely relevant to the repository horizon to begin with, but in any case, whatever you would have learned from those tests would no longer be relevant to a repository after the heat had dissolved and redistributed the silica around inside the mountain. So all the hydrologic measurements that you make at Busted Butte would not be applicable.

Another example are the niche tests. Those are very beautiful viewgraphs of all those tunnels in the mountain, and moving the water down and looking at the rates and the fracture flow and the pores. But once again, those tests are completely meaningless, because once you recognize the possibility that the silica can be redistributed by the heat and the water, all the hydrologic conclusions you draw from the way the rock behaves with the water under those ambient conditions are irrelevant to the way the repository is going to behave after the waste heat pulse rearranges it.

The third item, Jean commented about looking at sand as a backfill for the waste packages and doing some experiments to measure the interaction of the water and the heat and the sand. Those experiments have all been done a long time ago. There's a guy name Udell who's done a large number of those experiments, and I can tell you the answer after 20 or 30 days, the sand lithifies. The quartz sand dissolves and solidifies itself into a solid hunk.

There's a fundamental conceptual item that's missing from this program, and that is the idea that silica is mobile. It dissolves, it moves around, and it precipitates somewhere else, and that whole, that missing piece, that fundamental conceptual missing piece affects all the models and all the validations. It's a much more fundamental issue than whether the code is correct or whether the software is built correctly and whether the model that the software is representing is built correctly.

So on the admittedly longshot chance that my high school daughter's science project turns out to be correct and that the rock really does dissolve, I admit that skepticism is appropriate for that, this whole program has wasted very, very large number of millions of dollars doing, and is still doing, tests and analyses that either have already been done, the answers are obvious, or the results will be of no value to the program whatsoever.

I guess that's really all I have to say. Thanks.

COHON: Thank you, Mr. Matyskiela. Steve Frishman?

FRISHMAN: I'm Steve Frishman with the Nevada Agency for Nuclear Projects. I have two things. One is housekeeping, and that's with the Board's permission, I've asked Linda Lehman, who also is associated with our office, to take my place on the roundtable this afternoon because she was personally involved in INTRAVAL and I think she has some experience that is much more valuable for the Board to hear than anything that I might say about model validation in that context.

COHON: That's fine. Thanks.

FRISHMAN: The other is I understand that you still have not decided how you want to deal with the draft environmental impact statement that the Department of Energy has put out. And I think, just from the standpoint of my opinion, that you are going to have to deal with it, and I think it's important that you do, first of all, because you're a public advisory committee. And the public, this document is to, among other things, provide an avenue for the public to evaluate the project, evaluate within a context that is an accepted context for all major federal actions that have significant effect on the environment. And people are expected to comment on this if they have an interest, and I think it's within your charge as a public advisory committee to represent the public in this process.

And I'm not sure that the way you are constructed as an advisory committee means that you have to comment on all aspects of the environmental impact statement. I think it would be reasonable if you stayed within your statutory charge to evaluate the technical validity of the project, or the program.

And I also think that it's important because you're in essentially a unique position compared to the general public who is having to deal with this environment impact statement, and I think it's important that you have to bear the same burden that the public does, but you know a lot more, so you know exactly what that burden is. And that burden is that this environmental impact statement is to accompany a site recommendation, and you've spent at least the last day and a half, and much more out of your life, fully understanding that the project that is described and evaluated in the environmental impact statement for site recommendation is not the project that is the subject of site recommendation.

And it's become just in the last day and a half it's absolutely clear that the description of the project that the public has the burden of trying to comment on is not the project, the impacts are not the same. The impacts, despite what the EIS says, are not bounded for the design to be almost anything.

So I think while it may seem a burden to you to have to do it, I think your answer can be a pretty simple one, and I'm not going to try to dictate that answer, but it won't be very difficult to evaluate whether the Department did a pretty good job in evaluating the impacts of the proposed action, because the proposed action is not the same as what you know is going to be the proposed action in the site recommendation.

So I think the value that you can do in this public process, which is somewhat tortured, and I think once again I'll say the public is being imposed upon to spend whatever amount of effort and resource it can to comment on a document that essentially doesn't represent anything.

Now, I think it's important that you sort of, because of your special level of knowledge, take the lead for the public comment and make your understanding known without having to do very much digging at all. In the agency where I work, we're having to make a very major effort on something that I feel is a waste of our time and resources, because we're having to evaluate something that doesn't represent what its companion document, the site recommendation report, is going to talk about.

So I think you could probably help all of us who are the public, though some of us may be under different roofs of the public, I think you could help by at least reviewing the draft environmental impact statement according to your very special knowledge.

Thank you.

COHON: Steve, could I ask you a specific I guess legal question? If as you say there is a disconnect between what DOE eventually recommends and let's say the Secretary approves and the President approves, with the alternative in the EIS, doesn't that disconnect have to catch up with the process at some point?

FRISHMAN: It's supposed to, yes.

COHON: At least at licensing; right?

FRISHMAN: No, it's got to catch up in the NEPA process.

COHON: Okay. The final environmental impact statement is supposed to represent, among other things, a description of the project. And there are checks in this process that would--

FRISHMAN: Right. There are a number of ways that the Department could deal with the fact that the draft EIS doesn't represent what they even think the project is today. And there are means of doing that to come to a final environmental impact statement that in fact a sufficient statement.

COHON: I'm sorry. I meant checks that exist outside of DOE itself. I mean, would you have to intervene, for example, to make sure to make this point, or are there check points along the way?

FRISHMAN: The ultimate is legal intervention. The Department can avoid that, and they can avoid that if they get told by enough people that the final environmental impact statement must describe the proposed project, or the proposed action. And there are ways to get there from here, but if the proposed action in the final EIS is substantially different from that that was evaluated in the draft EIS, there's some procedures that have to be followed. And if those procedures aren't followed, then people are entitled to seek legal remedy.

And what I'm asking is that you use your special knowledge of the proposed action versus what is described in the draft EIS as the proposed action, to maybe encourage the Department to follow some procedures that will avoid the intervention, and also will in some way mean that the public didn't just totally waste its time reviewing something that they should not have been asked to spend their time and resources reviewing in the first place.

COHON: Got it.

FRISHMAN: I think that's where the service can be. You can use what you know to help make sure that ultimately, the process is one in which the public is genuinely involved.

COHON: Thank you. Judy Treichel?

TREICHEL: Judy Treichel, Nevada Nuclear Waste Task Force.

You know, even if I hadn't wanted to say something, after fitting your description of the unreasonable, fearful person, I would have to come up here, and I think that's really an important thing that Abe said earlier. People having reasonable assurance, reasonable expectations, but then suffering from an unreasonable fear of DOE, since I live in the west with other people who have previously been down-winders and probably still are. And part of that goes to the question that was asked yesterday by Dr. Sagüés when he was asking about possible health effect in the term that the public understands health effects to be, not the dead Nevadan, not the fatal cancer that wouldn't have occurred except for this problem, this project having been imposed upon the dose receptor.

But, yes, there is evidence and there's a lot of talk now about Beer 7 meeting to once again take up the question of low dose radiation exposure over long periods of time, and everybody doesn't just drop dead from the right cancer. There are generational things, and the fact that NRC yesterday was comfortable in being the person to leap to the microphone and saying no, we only deal with latent, fatal cancers, that brings about a fear, and I don't think it's unreasonable.

And in the case of Paul's airplane, he doesn't have to get on it. He never has to fly again if he develops a real fear of flying. And you're talking about people who are having a site forced on them. They are not consenting adults or dealing with informed consent in any way. Nevada is very, very much opposed to this project. And so the wording, the semantics become very important when you hear constantly that people have to be able to defend decisions, defensibility.

I know it's used one way by the people who work on the project, but it's heard in another way, and the kind of doing the best we can sorts of attitudes that you see here, because in the presentations that you see, there's always an effort to improve confidence, and it's usually DOE's own confidence. It doesn't seem to trickle down to the public that's having this project imposed upon them, and the enhancements that are brought up sort of are intended to rule out ruling out the project.

So one of the things that's wrong with the EIS, and that we complained heartily about all the way through, is that it didn't require them to state the need for the project. There was never to be a discussion about whether or not you needed a Yucca Mountain repository, and that's basic to everything here, because you're not going to get a willing public on a project that they don't see the need for, and to be expected to take a risk.

We're about to go into a discussion with the NRC very soon about risk communication and what kind of risk is reasonable and acceptable. Well, for the Yucca Mountain repository, no risk for Nevada, and it's not like, you know, you've used the analogy that your kid or your grandchild needs a kidney, and you happen to be a match, there's a risk involved there. But you would probably decide to do that because of the need, because of the benefit, you know, that you could certainly understand. But you don't take a risk for something like this.

And so all of the confidence, all of the validity, all of the--you know, I talk about them as possibilistic models because I don't see that a model tells you anything. I've got a file that I've started since this project called things that can't happen, and it's getting larger and larger and larger, and we've all seen those things.

So it's very important that you pay attention to this stuff and that you have courage and you really hit it hard, because the public, as the public representative, the public doesn't have any place to take its arguments. We can't go anywhere to say we don't like the idea that a health effect is a dead person. We've always come in too late for when such basic things have taken place, or when--you know, Nevadans weren't even on the scope when the decision was made for a geologic repository, and yet they have to be the ones that would accept this decision.

So we always seem to be kind of out of scope, or in front of the wrong audience, and an awful lot of these decisions are made by Congress, and we really don't have access. So we have to depend upon the courage of DOE investigators or the Technical Review Board or the NRC, and there's a tremendous lack of courage in some of those places. The Technical Review Board has been the best group that we have come across as far as inviting public opinion, making it easy for the public to play a part, and I really appreciate that, and many other people do, too. You get very high marks in Nevada.

But I wish there was a place where all of this could be laid out, and it's possible that it might be the focus groups that were mentioned, or the audiences that you mentioned to Abe.

Thank you.

COHON: Thank you, Judy.

I have a question following up on your comment. I don't know if it's for you to answer or for someone else. But with regard to the need for--wasn't it dealt with by Congress in the 1987 act?

TREICHEL: Oh, yes, sure, they gave them a free ride.

COHON: All right.

TREICHEL: Well, we can't go and talk about that.

COHON: I understand. That's just for clarification. Thank you.

Is there anybody else who cares to make a comment or wishes to ask a question at this time? This is the last public comment period, by the way. Yes, please identify yourself.

KONIKOW: I'm Leonard Konikow with the USGS. I'd like to ask Bob Andrews, based on his talk of model validation, with all the models and model validation exercises that have been done on the Yucca Mountain project for the last 15 years, what per cent of these exercise had led to invalidation of models?

COHON: You have to talk into a microphone, Bob.

ANDREWS: I'm not exactly sure, quite frankly. I think there were some earlier on in UZ flow that were determined to be invalid, if you will, back in the early Nineties, probably

'92, '93 time frame, that maybe Bo can talk to more than I. I'm not sure about the coupled process models, the thermal type models in the drift. I'm not sure whether any of those were determined to be invalid. I think they reasonably matched.

I'm not sure if there were other ones that were invalidated. The only one I can think of right off the top of my head, quite frankly, is the UZ flow model back in the early Nineties was invalid.

CRAIG: What happened to the old saturated zone model?

ANDREWS: Oh, okay, yeah, that's a good one. The saturated zone flow model done prior to VA at the site scale was determined to be invalid because of flow directions, of course there's limited data also, but the prevalent view was that flow model was invalid for how the flow system was characterized south of the site. So it was not used, in fact, in the VA because of that, and a more simplified representation was chosen instead.

So those are the two examples of invalidity, but I think it's a worthwhile--it's a good question, and we'll probably bring that up later on this afternoon with the examples from Bo and Joe, too.

KONIKOW: Well, hopefully on this roundtable discussion this afternoon, I'll have an opportunity to give you some details of why I think the whole concept of validation as you do it is misguided and probably damaging to your own cause, and so we'll leave that for this afternoon.

COHON: I couldn't ask for a better preview for this afternoon's meeting. What a great teaser. I'm sure the afternoon will prove as interesting, at least as interesting and enjoyable and enlightening as the morning has.

Thank you again to our morning speakers. We stand adjourned now until 1 o'clock.

(Whereupon, the lunch recess was taken.)

AFTERNOON SESSION

CRAIG: Okay, this afternoon, we have the first part of the afternoon prior to the break, main break in any event, we have two talks. The first one is unsaturated zone model validation by Bo Bodvarsson from LBL, and then he will be followed by Joe Farmer from Livermore. And I am happy to note that this is an all Berkeley crowd. Bo's Ph.D. is from UC Berkeley in hydrogeology, and Joe Farmer's is from Berkeley in chemistry. But we begin with Bo.

BODVARSSON: Okay, can everybody that wants to hear me hear me?

My name is Bo Bodvarsson. I'm going to talk a little bit about the unsaturated zone model validation and the repository safety strategy.

My talk, this is the outline of my talk, and I'm going to put it here on the right so you can always look and make sure where I am with the talk. I'm going to talk a little bit about what the UZ flow and transport model is, how it relates to the principal factors, and development of the UZ model that's been going on for a decade or so, calibration of it, a little bit about the use of the model, uncertainties of the UZ model, then validation of the UZ model.

I got a request real late, about a week ago, from the Board asking that I talk about seepage. That was not really my intent here, but I have a few viewgraphs in the end talking about the latest calibration seepage model, and any questions that you have, I'll be glad to answer about any of these models.

So what is the unsaturated flow and transport model? It's very simple. It basically computes the flow of water, of chemicals and heat and gas throughout the mountain, anywhere in the mountain.

So the main processes you see here on the left-hand side, of course you have infiltration coming into the mountain that vary spatially. You have water flowing through the fractures and the matrix block, and the fracture/matrix interaction is a key problem. You have seepage into drifts. Some of the infiltrating water will seep into the drifts, a small amount hopefully. We have complications due to perched water. That has been one of the most important data sets that we use for calibration. And then of course we have to quantify sorption in the Calico Hills. That means how much of the radionuclides that go from the repository are actually sorbed and don't go into the saturated zone. And here are little schematics showing fracture/matrix interaction, infiltration and the waste package.

Now, the UZ flow and transport model and the UZ flow and transport PMR consists of roughly six models. Always think models. I listed four of the most important ones, because those feed performance assessment, and that is the properties model, that is the model that determines permeability, porosity, as van Knuckten talked to, or anything else that deals with flow of water and gas and chemicals and heat. We have then the flow and transport model. This is the three dimensional representation of flow patterns in the mountain. We have the seepage model that quantifies the amount of water seeping into the drifts. And we have the thermohydrologic chemical model on the drift scale that basically changes and modifies permeabilities and porosities because of precipitation and dissolution of minerals due to heat and coupled effects.

Those are the four models. And then we started this process of deciding what to talk to in this talk. I picked the flow and transport model. I could have picked any one of these four models, and I just picked that one because that has a reasonable amount of calibration data, as well as validation exercises.

I will then also talk a little bit about the seepage model at your request.

Now, principal factors that feed this group of models is seepage into drifts and UZ sorption and matrix diffusion, as you're well aware of. Then we have some seven other factors that are directly related to the UZ flow and transport PMR.

Now, very briefly to tell you about the data, because a model is no good without data, although nobody can prove you wrong if you don't have any data. Fortunately, we have quite a lot of information from the mountain. We have the gas pressures that has been extremely useful to determine the permeability structure everywhere in the mountain, because these signals, even though they are tiny and you can just barely feel them, we monitor them all throughout the mountain.

We have then of course saturation and water potentials from cores. We have a bunch of tritium, Carbon-14 and geochemistry, including total chlorides and sulfides and Chloride-36, and all of those, which are proven to be very, very useful. We has gas data and ages of gases incurred from Carbon-14, and young gases shallow and old gases deep, and we have of course temperature data that helps with the percolation flux, and we have a lot of ESF data and east-west cross-drift data that we use.

Now, why do we do a UZ flow and transport model? Why is it needed? Number 1, you need to integrate all of this data into a computational framework. A sole type distribution in a mountain doesn't tell you anything, but when you compute it with a model and match it, it tells you something about the amount of flow and the flow patterns.

You also want to quantify the water, gas, tracer/radionuclides and heat transport in the UZ under various assumptions by varying conceptual models, by looking at different parameter distributions, basically looking and varying things that we consider uncertain in the mountain, and getting the distributions of flow patterns, groundwater travel times, and things of that sort.

And, of course, we want to provide this calibrated UZ flow model to PA for their TSPA calculations.

This is a very, very simple generic logic diagram, and Priscilla and Bob Andrews were talking something about this this morning, and it has to do with calibration, field data, predictions, comparisons, validations, and this is my simple mind at work here. You take--let's take a process such as gas flow in the mountain, and let's say we have a signal on the surface and we have sensors below, and we predict, we take the field data and we stick it in the model and we predict the pressure variation of all the sensors in the mountain. That's the test. That's a test.

We then compare these predictions and observations, and actually in this case, we did this over many years, where they did not send us their data set, they sent us the surface pressures, they kept the data set, and until we sent them our results, it was really a blind mass. And then you compare predictions to observations, and if they're acceptable, and I don't know how to define acceptable--Bob Andrews knows how to do that--so if they are acceptable, you go down here and you say my model is calibrated for this process at that scale, and can, therefore, be used for that process on that scale. If it is not, we go at it again. We recalibrate, we get more field data. Of course the prediction data is always different from the calibration data.

So I'm going to show you now--talk a little bit about the development, and I'm going to talk a little bit about the calibrations to give you some confidence in this model that's reasonable, and we will start with the pneumatic data that we just talked about.

We have it available for quite a lot of boreholes. We use it to estimate large scale fracture and fault diffusivities, and we get those, fracture and fault permeabilities is what we get out of this. And you see here you can have it distinguish between the simulations and observations, because the models predict really well what's going on. Here, this doesn't show it very well, sorry about that, what happens here is that you see the ESF hitting a fault close to this borehole, NRG-7a, and because of that, the signal changed because it short-circuited through the ESF into the fault, and laterally through the fault. So you see much more variability in the signal here because it short-circuited through the ESF.

Now, what does that give us? That gives us directly permeabilities of that fault along this lateral pathway.

Then you have signal and many sensors here. Of course the more amplitude, then the more, or the higher the amplitude, the closer to the surface, this is Tiva here, then you go into TPM, and then you go into Topopah. And, again, the model matches very well the data.

Feel free to ask questions during this if you want to, or is it a rule you can't do that? I don't know.

Another thing that we compared to is the saturation and moisture data, and we frequently when we show this data set, people say, I mean they don't have a clue what you're doing here, because it goes apparently all over the board.

This is the nature of water potentials. Water potential is very hard to accurately measure. They are plus or minus a bar. Therefore, we do not expect to match this, because the data errors are that much.

Saturations are much more easy to measure because you take a core, you weigh it, you dry it, you weigh it again, and you get saturation. So we match that there for most of these boreholes.

I remember a question that I guess the distinguished Chairman asked a couple of years ago, and says what makes you think this is a good match, and that's a very good question. What we do is we simultaneously match all eleven boreholes, every one of them we simultaneously match with the ICOP code. We do this statistically so we get statistical maps, give them the input volumes. For example, we can weigh each saturation point ten times more than each water potential point if we believe this data is more reasonable.

Therefore, for each borehole, we are not going to get an exact match because we are matching all of them simultaneously. But on the average, you get the layer properties, a very good indication of layer properties as well as all the statistics that go with it, the variability between boreholes, and things like that.

This is a very interesting data set that we just started to work on recently and, therefore, this is work in progress, but I wanted to show it to you because we always want to update the best we can. This is data from June Fabryka-Martin and Al Yang of USGS, June from Los Alamos. This shows here the east-west cross-drift results. They show the chloride data here in one of these triangles, and what they show here is our prediction of the chloride data before the ECRB. This is based on Alan Flint's infiltration maps, and you see here we have much too high chloride values here, and we have much too low here.

Now, chloride relates directly to infiltration. The higher the chloride, the less infiltration. The lower the chloride, the more infiltration. Just simply you have a fixed source of chloride at the surface, and the more water you add to it, the more you follow the chlorides. It's as simple as that.

We used this to now do an exercise, and remember it didn't match very well, so we can't say that our model is validated against chloride, can we? So we went back to calibrate, and we changed the infiltrate map, because I believe the infiltration map is the reason for this error. The chloride source is very well known and, therefore, this should be a very good indication of the percolation flux or infiltration flux.

BULLEN: Bo, this is Bullen, Board. You asked for this, and so you're going to get the question.

Isn't the movement of the chloride also going to be associated with lateral diversion in the UZ zone above?

BODVARSSON: Yes.

BULLEN: So the data that you got from June Fabryka-Martin here could have been smeared or smushed out because of the fact that you've moved it from where there would have been a high infiltration rate, to where it actually came down fractures, or whatever pathway it came in?

BODVARSSON: Yes.

BULLEN: And so does that pose a big difficulty in calibrating then when you have that kind of lateral diversion?

BODVARSSON: No, because the 3-D model, they use the full 3-D model to calibrate, and it doesn't mean, and you're right that I can say that within a hundred meter interval, make sure that this chloride signal is exactly there. You're absolutely right. But you have a lot of capillary equilibrium, you have diversion due to capillary pressure, and things like that. You're absolutely right.

But when you look at the data set here, it's very similar values for this data set. And this is actually the map we obtain by assuming just a single value for infiltration. Therefore, very low variability, and I'm going to show you that next.

BULLEN: Okay.

BODVARSSON: This is the infiltration map, and I think this in some sense is really good news, if this is right. Why is that? First of all, we don't have the high infiltration at the crest that the infiltration models say 20 millimeters per year, 30 millimeters per year, up to 60 millimeters per year. The chloride says it varies between 10 milligrams per liter to 50 milligrams per liter. That corresponds to a flux of between 3 and 9. So I just said I want to make that 6, because I don't believe this variability, I don't believe six to eight and four are the same number. Right, Bob?

So that's really good news, I think. Now, why do I believe it? I believe it for one reason, one important reason, at least for myself. A long time ago, Ed Weeks told me I don't believe in high infiltration fluxes at the crest of the mountain because to me, the Tiva Canyon is very tight. There's nothing going to go in there. It's all going to run off. This is exactly what we are seeing, the same rainfall, but it all gets run off down the mountain. It makes sense. Gravity kind of wants things to go down.

Then it also makes sense when you look at these areas, that basically the high elevations here where you expect more rainfall, you get more infiltration. The thick alluvium areas, you have almost no infiltration, and then in between, you have the runoffs and the rainfall in the intermediate areas.

The data we used to match this is all on the ESF data from June, all of the east-west cross-drift data from June, all of the borehole data.

NELSON: Has there been any indication that there's any infiltration coming in from the Solitario Canyon itself?

BODVARSSON: That's a very good question. A year ago, I would have said exactly that is a very good case for that because we used to believe we had inversions in 14-Hs and in borehole ST-9 and ST-12. The survey has since changed their mind and said that there's not an inversion, that maybe there's purely vertical flow there. So right now, we don't have sufficient data, Priscilla, to say if there is a lot more there.

NELSON: This is Nelson, Board, again. Is it important to know the answer to that?

BODVARSSON: Yes. It's very important to know the answer, and the reason is this. We talk a lot about pulses. We talk a lot about rainfall infiltration occurs once every five years through two days, four days, whatever. In the middle of the repository, what is happening is here's the repository area. We have PTN on top of the repository area everywhere except close to the Solitario Canyon. PTN is what diffuses pulses, because it's a porous medium, 40 per cent porosity, 300 millidarcies permeability. It doesn't allow anything through it in less than 500 to 1000 years, and doesn't allow these pulses to occur except close to the fault, like June Fabryka-Martin shows.

Now, here close to Solitario Canyon, we don't have that. It's exposed, and you get infiltration directly into the Topopah Springs Unit. You have very fast fracture point in the Topopah Springs Unit, and you might get, if there is thick infiltration there, you might get significant seepage in that area. So we need to look at the pulses in that area.

NELSON: Nelson, Board. Just one last thing.

It seems like the yellow area is bounded by, I suppose it could be topography, but also by faults.

BODVARSSON: Yeah.

NELSON: To what extent is the fault presence dominating infiltration?

BODVARSSON: That's a very good question. But the honest answer, Priscilla, is that that's just how we drew it. We really don't know. I have data points coming here, and I know that it's about six years. I have no idea how to do this area here, because I don't have any boreholes in this area here. So I just said my yellow is this, and I made it so that it corresponds to a fault.

BULLEN: Bullen, Board. You actually just raised something that goes back to confirmatory testing, which is well beyond site recommendation and licensing. But as you gain data, during the operational phase if we so choose to build a repository, do you expect this map to become much more detailed and more significant, and then we'll be able to continue to calibrate and update the performance models for closure?

BODVARSSON: Yes.

BULLEN: So I guess the expectation is that when you're at the horizon and you've got the data, because you've got the nice little data points on the ECRB and ESF, you'll have basically a nice map of what you expect the infiltration to be?

BODVARSSON: Yeah, except that--you can do that, I can go back and I can match all my ups and downs in my chlorides. I can do that. Now, is it worthwhile to do? No, because it doesn't make any difference, because I get between 3 and 9 millimeters per year, and that just doesn't have serious impact on seepage, nor on transport. So, therefore, these details won't matter.

BULLEN: Okay. Bullen, Board, again. The follow-on question then would be when you finally do climate change, will you expect to see some significant changes in your model if the infiltration rate at the top of the mountain goes to 140 millimeters a year?

BODVARSSON: Definitely.

BULLEN: So that's where you'd see the change?

BODVARSSON: Yes.

BULLEN: Okay.

PARIZEK: Parizek, Board. That yellow is not entirely arbitrary. The PTN is there, plus your high elevation; right? It's not anybody could have done that? You're saying, no, I'm using my geological map and elevation to decide on where the yellow border is?

BODVARSSON: See, I have is I have the ESF data here, so I have data along all of this thing here. I have data along all of this cross-drift. I have SD-9, I have SD-6, I have SD-7 here at the boundary, and that defines for me this region all here, all of this region pretty much is very easy to say here is six. And then the rest of it is more arbitrary. So it's not totally arbitrary at all. You have quite a lot of information.

PARIZEK: Yes, but I mean that tail to the south is along the ridge.

BODVARSSON: Yeah, the tail to the south is along the ridge. Yes. So that is purely hypothesis.

PARIZEK: Yes, that's a concept. You're carrying a conceptual understanding of it south.

BODVARSSON: That's exactly right. Using these basic ideas, we believe infiltration is related to the geological features and thickness of the alluvium and all of those.

Then we talk about perched water calibration. Like I said, perched water has tremendous effects on the calibration. It's extremely important. Why is that? A, because we know pretty much the extent of the perched water from testing. B, we know the ages for Carbon-14. C, we know the chloride content and the chemistry, so it gives us tremendous information.

This is one conceptual model for perched water. One problem of the perched water is that even though we have significant effects on dilution, matrix diffusion and sorption, just because of what the bore tests brought up over the last couple of days, that is, the distribution of zeolitic rocks and vitric rock in the Calico Hills makes a difference in sorption.

It's obviously, for example, when neptunium sorption in zeolites is poor, sorption in vitric is one. If it is more than one, sorption means a heck of a lot. So we are right now carrying three conceptual models on perched water through to PA to look at the sensitivity of this important conceptual model for PA, for SR.

This is predictions of Chloride-36 and also for strontium. Strontium is a very strong indicator of the presence of zeolites, because strontium exchanges and sorbs through the zeolites. So you see here a drastic reduction in the strontium content in these boreholes due to the presence of zeolitic rocks in the Calico Hills and Prow Pass.

Also, strontium is very much related to infiltration and percolation flux. We are going to use these data here to compare to our map, we just got the map last week, to make sure that this is consistent with our now current idea in progress about infiltration.

The Chloride-36 I've always found to be much less important. We talk a lot about it, but what does it do for us? I believe there's every indication and all the data suggests very strongly that this is a very minor part of the flow, much less than 1 per cent.

Now, I'm going to go into uncertainties. I want to say a few words about the use of a UZ model and then I'm going to go into uncertainties.

As you know, the model is primarily used by Bob Andrews and his group. We just finished calculating 30 three dimensional flow fields based on various assumptions and conceptual models that we are in the process of transferring to PA for them to start their base case calculations of TSPA for SR Rev. 0. So that's enough about the use, I guess.

I want to talk a little bit about the uncertainties of the model, and of the data, and this is just my notion. This is just my idea when I look at the model development over the last few years, where we are going to be at site recommendation.

These are uncertainties. They vary tremendously in importance. Some of them are much more important than others. We have infiltration, water properties, fracture and fault properties, all the way down to detailed flow mechanisms.

These are the plans to address them that Jean Younker and Mark Peters mentioned in their presentations, and I'll just walk you very, very quickly through this.

Infiltration and future climate we are now starting--to use all the chemistry and temperature to integrate it in the infiltration model that we hope will be more reliable than what we have now.

Water properties from pneumatic tests, I think this will be--we have used the pneumatic test, fracture properties for our seepage models, for Alcove 1 models, and they seem to work just fine, and we're going to verify that, so I think the parameters can be very low by SR. We have confidence in this.

Fracture and fault properties and variability. The fracture properties from pneumatics are very well handled. The fault properties of liquid flow is something that we need to look at.

Fracture/matrix interaction, we are using geochemical data like the chlorides and like strontium and others to model Alcove 1 data, Drift to Drift data, Busted Butte data and other geochemical data to validate what we call the active fracture model, which is a model we just published in Water Resources Research about a year and a half or two years ago that says depending on the infiltration rate, only a small fracture of the total fractures in the mountain flow. The more you put in, the more fracture flows. And we are using that i all of our UZ models as well as all of the PA models that follow the UZ model. If you want, I can send you a preprint of this article.

Fracture and matrix sorption. We are not relying on fracture sorptions right now. We are relying on matrix sorption. We use Busted Butte data to validate laboratory measurements of sorption in the vitric Calico Hills. Busted Butte has very limited zeolitic Calico Hills, so we can only us it for the vitric part of the Calico Hills.

I'm going to say a little bit more about that in the validation exercise that's coming up.

Colloidal transport, we are using LANL. Los Alamos is using laboratory data and analog data to do a colloidal model, and right now, we don't have much confidence, but I think that will be medium by the time of SR.

Thermal effects on flow and transport, also detailed flow mechanism. I believe it's very, very difficult for us to determine exactly where the flow paths are, how far between they are, and things of that sort, so this is difficult for us to evaluate.

Now I'm going to talk about some validation examples. We've gone through the calibration and we've gone through some of the uncertainties, and now we're going to talk about validation and I'm going to give you some examples here.

The first one is pneumatic again. Again, like I told you, we have blind predictions that we do with the pneumatic, and they give excellent matches with all sensors after calibration. So I believe that our gas flow components of the UZ model are pretty well validated on this scale.

This is Alcove 1, and Mark Peters and Jean talked a little bit about Alcove 1. This has proven to be an extremely interesting and good exercise for two reasons. One is seepage and the other one is matrix diffusion.

Seepage, even though we put thousands and thousands of millimeters per year into Alcove 1, and I'm not going to go into detail, only 10 per cent of it seeps. It's a low number, given the high percolation flux number. And this again verifies some of our model results. This is what we did. Here is the calibration activity with the flow in Phase 1. We then used that to predict Phase II flow, which is shown here in the blue. You can't even see that, but it's supposed to be blue. The red is the data; blue is the predictions here.

And then we also predicted tracer breakthrough. And this is the most important thing. This is the tracer breakthrough. This occurs without matrix diffusion. These occur with matrix diffusion, and the proper diffusion coefficient for bromide. That's basically the tracer we use.

Data points from the field are right here, just these three data points right here. So what you're seeing is not a lot of data you see, but the important thing is we only saw tracer breakthrough after some I think it was 30 or 40 days or so, and that's exactly what it says that matrix diffusion does.

So matrix diffusion is extremely efficient here. We estimate that half of the fractures between the surface and the alcove flow, and the matrix diffusion is very efficient in retarding the tracer going through the mountain.

This is prediction for one borehole. This happens to be SD-6, which is the latest drilled borehole. For all of the boreholes that we are drilling, plus of course the east-west cross-drift, we predict before we drill the boreholes and before the east-west cross-drift. This shows some of the saturation data from this borehole, and we under estimate in this borehole the thickness of the Calico Hills vitric in the geological framework model. Other than that, it matches pretty well both the moisture tension and saturation.

This is Busted Butte data. This is Phase 1A, and if you remember from Busted Butte, there was an injection borehole for six months, and that was very, very slow gradual injection to mimic the flow through the mountain, and this is the extent of the measurement after they are recorded. And you see that there's about two meters or three meters and it spreads out a little bit here at the bottom. This is the model calculation that shows very similar spreading of this. This is the tracers. We don't have tracer measurements yet from this, so it's very similar shape from the model prediction as this.

Now, there's several things I want to say about Busted Butte. A, Busted Butte is only the vitric part of the Calico Hills, not the zeolitic part of the Calico Hills. B, the vitric part of the Calico Hills is porous medium, no fractures. Whatever fractures are in there are immaterial because the permeability of this stuff is a darcy. So fractures are not fractures that seep back into the matrix. So fractures are immaterial here. C, it follows exactly the capillary pressure theory that we are using in the models and have been using in the models over the last five or ten years. The extent of this data set is matched equally well with the 1997 viability data set from the UZ model.

What's the differences? The difference is viability data set, has permeabilities on the order of 100 millidarcies. The Busted Butte data is about 1000 millidarcies. So far, all of the data I've seen for Busted Butte verifies what we are using in the models in terms of flow mechanism and sorptions. That means there's nothing to transfer from Busted Butte to the Yucca Mountain right now because it's immaterial. We are not conquering anything. We are matching what is right there, and what we have measured for Yucca Mountain.

SAGÜÉS: Excuse me. This is Sagüés. I don't know exactly if the picture at the bottom is the same scale as the--

BODVARSSON: Yes, it's the same scale.

SAGÜÉS: And what is the meaning in the picture in the bottom? Where's the meaning of the colors?

BODVARSSON: Well, this is a fluorescein type of thing.

SAGÜÉS: And the boundary of that oval like region in there corresponds to what kind of concentration? In other words, is it directly comparable to the picture above, or is it just simply a coincidence that it happens to look the same?

BODVARSSON: We do not have at this point measurements in concentration as a function of space in this. So I cannot compare my concentration to this one here. But what I'm trying to say, all the parameters and all the models we have been using over the last five years are not extremely sensitive to anything but capillary suction, which is why this spreads out. You don't see much of a gravity component here. The infiltration rate is so small it just spreads out like that, due to the capillary functions that we use for the vitric Calico Hills that comes from measurements from Lorrie Flint on the actual vitric Calico Hills.

SAGÜÉS: What kind of a spread would you have seen if capillary action wouldn't have been the main element? What would it have looked like?

BODVARSSON: Vertical. You see, we are not doing an analytical solution of this. What you will see is regardless of the parameters, you can, in dimensional space, you have a point source. It's going to develop by halo, and depending on the properties, the halo, how far up it goes and all of that, the stronger the capillary function is, the more the vertical drive of the fluid obviously. The smaller it is, the less. And if there is no capillary function, you just have gravity flow.

SAGÜÉS: So really, what I'm trying to say is the pictures sort of look vaguely similar. But you will expect if you just put ink in the center of paper, it will spread out in all directions. But, I mean, the picture down there sort of vaguely resembles the one at the top. It doesn't have any particular quantitative meaning at this time; is that correct?

BODVARSSON: Well, it has a lot of meaning to me for the following reasons. Your flow from the repository through the water table occurs through the Topopah Springs into the Calico Hills vitric or zeolitic, and out through the water table. Flow through the Topopah Spring is a fracture dominated flow. Therefore, the source term going from Topopah Spring into the vitric Calico Hills, where we are taking credit for sorption, is going to be a point source in space that varies. It's not like a porous medium. There's a spacing of some ten meters, twenty meters, we don't know yet.

Now, the fact that the Busted Butte data show this strong capillary spreading of this indicates strongly to me that this point source is going to spread a lot in the Calico Hills, and we can take full credit for sorption over the entire Calico Hills.

SAGÜÉS: Sure. But that's a qualitative--

BODVARSSON: That's a qualitative solution. We can never make this qualitatively. That's why I didn't spend a lot of time to make this exactly the same as this when we don't have the tracer concentrations. We are waiting for the tracer concentration to make a definite--

SAGÜÉS: Right.

BULLEN: Bullen, Board. Before you leave that one, does that mean that source term when you're coming out of the Calico Hills is then a planar source?

BODVARSSON: Yes.

BULLEN: Okay.

BODVARSSON: No, no, hold on.

BULLEN: What causes it to come out then, is the question.

BODVARSSON: Well, it's a good question. We have two areas in the Calico Hills, and your questions about the Calico Hills are very good. We have the northern area, which is zeolitic, and we have the southern area which is vitric. The vitric part of the Calico Hills is a porous medium, just like you said, and will spread all out, and you will have a planar source at the bottom. But below the Calico Hills vitric, there is Prow Pass zeolitic, which is again low permeability to fractures. Flow is going to go out of the vitric either into that or as a perched water down that through the water table. We don't know exactly.

Does that answer your question?

BULLEN: Yes. Thank you.

BODVARSSON: In the northern part, we have more problem with the zeolitic. That's this conceptual model for perched water. One conceptual model is simply nothing goes through the zeolite, and right now, we don't take any credit in PA for sorption in the zeolite because of the possibility of lateral flow down the faults.

The other conceptual model that we're looking at now trying to take credit for the zeolitic rock is vertical flow, and we're looking at the chemistry through there.

This is cross-drift calculation. This is percolation flux based on Alan Flint, and this is strontium variability in the east-west cross-drift. And I just show it to you to show that we actually predict a lot of stuff for the cross-drift. Right now, we don't have any information to verify this yet.

NELSON: Can you tell me again what that plot is? Because I was trying to see it.

BODVARSSON: This one here?

NELSON: Yeah.

BODVARSSON: This is strontium, three dimensional use, same as the chloride, we now put strontium on the surface in the infiltrating water, and Brian Marshal is in the audience, does a lot of work on strontium, and we predict what the variability in strontium would be in a cross-section, including the east-west cross-drift. I mean, I want to make measurements of strontium and compare it to see if we have accurately predicted this.

I'm almost finished. I was asked, this is not of my own doing, I was asked to provide an external peer review list, and here it is. We have been reviewed to death almost.

Before going to seepage, I just want to summarize this part. I feel the UZ model is reasonably well calibrated because nobody can define reasonably well, so that should be okay against all available data.

Uncertainties vary significantly in the different components of the model. Some, such as gas flow, are very well understood. Others, such as matrix diffusion, are less understood.

Current field activities should certainly increase confidence and reduce uncertainties.

Model calibration and validation activities yield confidence in model predictions of some processes, such as gas flow, bulk water flow and transport through the Calico

Hills vitric. And I don't see zeolitic here.

Less data are available for calibration and validation of other important processes that we must concentrate on, such as matrix diffusion and transport through the Calico Hills zeolitic.

The UZ model uncertainty will continue to decrease due to additional calibrations and validations using Yucca Mountain and natural analog data.

So that's enough for that, and I can do seepage real quick.

NELSON: Can I just ask you a question?

BODVARSSON: Yes.

NELSON: I recall an observation that was reported on--this is Nelson, Board--about construction water penetration, and how much further it went in the non-lith as opposed to the lith. Would that have been predicted by--I mean, this is sort of leading towards the continuum treatment of the mountain here, so it wouldn't really work for the treatment of the equivalent continuum. But would that have been an anticipated information there that--

BODVARSSON: That's a good question, and I will try to answer it. I haven't thought a lot about it. I think the answer is probably no, and I think this model should predict it, because that's the purpose of this model, even though it's a continuum model, it still is a dual continuum with fracture flow and matrix flow, so we should be able to predict migration of fluids down through the mountain.

Now, the reason I say that probably--would probably not do it is because of two things. One is that we don't have very much hydrological data from the lower lithophysal, unfortunately, and most or all of it is from vertical boreholes. That's why Jean and Mark Peters said we are emphasizing systematic hydrological testing of the lower lithophysal to really get at that.

The second reason is just my own, because when I walk through this cross-drift, I see so totally different rock from the middle and lower lithophysal, at least in my mind, and I was personally surprised when I saw it.

Now, my geologists here, like Mark Tynan, may say that there's no surprise, but I was surprised. So the answer to your question is a good one, I think we would not have predicted it.

NELSON: Nelson, Board. And you're going to see that same difference in phenomenon in the percolation test, the seepage test, between the ECRB and the ESF, because of the two kinds of rocks that are present in the flow paths.

BODVARSSON: Yes, I couldn't agree more. I think we understand seepage in the middle and non-lithophysal, like I'll show you a little bit--I think we understand it quite well, but I couldn't tell you anything about lower lithophysal because I don't know how that different rock in my mind is going to behave.

SAGÜÉS: Before we go on to the next, on your Slide 13, which shows the UZ model calibration with Chlorine-36?

BODVARSSON: I should never have invited them to ask these questions.

SAGÜÉS: There it is. Is that along the ESF, that particular cross-section that you're showing there?

BODVARSSON: Yes, the ESF.

SAGÜÉS: That's the ESF. And you're getting the elevation information from the different boreholes; right? Like, for example, I see there that there is the WT-2. Is that the borehole?

BODVARSSON: Yeah, WT-2 is a borehole.

SAGÜÉS: Is a borehole. And then you SD-12 next to it. But in between those two boreholes, you have an orange region and a yellow one, with this little green thing in between. That resolution comes from--this is along the lines of the question that Priscilla was asking yesterday. Why is there so much fine detail in between what appears to be just simply--

BODVARSSON: It's because Alan Flint measures infiltration so precisely. Alan Flint, in this case here, this was--we used Alan Flint's infiltration map that has a 30 meter spacing on infiltration data. We input it into the three dimensional model, and that gives you the variability in all of the chemicals moving in through the mountain.

SAGÜÉS: And the infiltration is measured what, at the surface?

BODVARSSON: The infiltration is measured at the surface. He believes there is a big difference between infiltration at ridge tops and in the crest of these little valleys and at the bottom where you have the thick alluvium. And that's reflected in great variability over a 100 meter distance.

SAGÜÉS: Are we going to have for something like this, are we as reviewers going to have something that says okay, in constructing this map, the following inputs were used?

BODVARSSON: Yes.

SAGÜÉS: This is borehole data, surface infiltration data. Okay, these are the inputs and this is the output. Because when I see this map, somehow there is a lot more input that, or maybe more input than what appears to me.

BODVARSSON: If the Board got the UZ model for the viability assessment, which has some 24 chapters in it, this happens to be Chapter 18, if I remember correctly, and that tells you all the details, what went in, what came out, for the Chlorine-36. And I assume the Board would have had that a long time ago. Is that right?

Any other questions?

BULLEN: Excuse me. Bullen, Board. Since you're on this viewgraph, this is the one I was going to ask my question on anyway, you make the statement under the second bullet that bombpulse Chlorine-36 indicates the presence of fast paths, and currently believed to constitute less than a per cent of the flow. That's a very important statement. And can you tell me the basis for it, and the experiments that you might want to do that would bolster your confidence in that it's less than 1 per cent of the flow?

BODVARSSON: Okay. Well, number one, I will put a caveat on this now from the start. For example, we never know how much flow goes through each flow path according to Chlorine-36. Chlorine-36 just says it got there. It doesn't know how much it is.

But the reason I believe it strongly in my mind, and I should have put this is what I believe, is the following. We have done a bunch of measurements of Chloride-36 trying to look for Chloride-36 measured much, much more close to fault that anywhere else systematically in the mountain. And even though we looked and looked and looked and looked, the ratio of bombpulse to non-bombpulse, Chloride-36 is much less than one, even though we looked and looked and looked.

BULLEN: Okay. Bullen, Board, again. Is there any experiments that you're planning on doing in any of these things that will help you further define the fact that it's 1 per cent or less than 1 per cent? Or are you just going to have to use the measurements that you've got as the basis for that conclusion?

BODVARSSON: Well, we did use the cross-drift. We predicted, June Fabryka-Martin predicted the east-west cross-drift. You will find it in two locations and two locations only. We found it in two locations and two locations only. And then I'll use the rest of that to try to verify this, but I don't know of any other.

BULLEN: Thank you.

COHON: Cohon, Board. Could you put up Slide 7?

You made quite an understandable observation about the word "acceptable" and how difficult it is to estimate that or to arrive at that. Who decides whether it's acceptable? Is that your decision?

BODVARSSON: Well, I think it's a joint decision by PA and the process model developer, which is me. Basically, what I believe is that the word "acceptable" is not so hard to do, and the reason is the following. I believe you need to put emphasis, and Bob said this already, you need to put emphasis in validation of where that model and what scale is going to be used for in performance assessment. Okay?

Therefore, when you take a look at, for example, matrix diffusion, I showed you Alcove 1, we can look at that and decide in our minds based on impact from PA, if the uncertainties in the parameters we get from matrix diffusion significantly affect PA or not. If they do not, that is acceptable to me. But if they do, it's not acceptable.

COHON: Will there be quantitative criteria to arrive at acceptability, or will it be purely qualitative?

BODVARSSON: Maybe I should ask the higher ups. I think it will be qualitative, personally. I think we will--well, maybe I shouldn't say anything. Maybe the best thing to say is say nothing.

COHON: Well, Abe is nodding his head, so I guess you're right.

BODVARSSON: Okay. I'll say that then.

COHON: Second--one more question. You make a clear distinction in this diagram between calibration on the one hand and validation on the other.

BODVARSSON: Yes.

COHON: And in your summary, I couldn't help but notice that while you said the UZ model is reasonably well calibrated, you said nothing about its validation.

BODVARSSON: Yes.

COHON: Do you want to say something about its validation?

BODVARSSON: Yes. There were some words in there that didn't mention validation, but what I mean to say is that I think for some processes, it's already validated, like gas flow processes on a mountain scale. Because we have so much data and every data, we calibrate it very well, we predict it very well, and things like that. All the processes, like matrix diffusion, we have very low data, it's not validated.

COHON: Has PA agreed with you on those claims of validation?

BODVARSSON: I think so. I think so.

COHON: Bob Andrews is nodding his head.

BODVARSSON: All right.

COHON: And so is he. Thank you.

PARIZEK: Parizek, Board. You mentioned neptunium, you would have a value of four in the non-vitric part, and it would be one in the vitric.

BODVARSSON: Yes.

PARIZEK: If you're not sure whether vitric or non-vitric exists down there, what do you do, put one? Or did you put a one and two and a three and a four?

BODVARSSON: No. See, I believe we know a heck of a lot more about where the vitric is than perhaps the Board does. And I can give you a reason for that.

For example, you have H-5. H-5 is the first bore identifies the thick vitric, or vitric zone in the Calico Hills. We didn't find the zeolitic rocks up north. We found the vitric on the south. SD-6, we just drilled, Mark Tyner and I actually located that borehole to find out the extent of this hole in the zeolitic rock in the vitric part, and I went as far north as I dared to go to try to make sure that I would find vitric there, and that's where the vitric is.

In our PA calculations, we have a conservative volume for the vitric part we are taking credit for, and we are not taking credit for the zeolitic rocks.

So basically, I would say that there might be more potential than we are using, because we are being very conservative because of the limited data.

PARIZEK: That would be the case you have to make for NRC, as an example?

BODVARSSON: I you want to take more credit, you would have to get additional data and take more credit.

PARIZEK: The Figure 14 showed some use of chemical data, and it seemed like much of that was for tracer value showing this mass of water did in fact go through the rock, or was that to deal with chemical interactions, such as--this is on Figure 14, you had a discussion about the use of chemistry, putting more chemistry data into your models.

BODVARSSON: No, the chemistry model, I think we are on the right track getting better percolation values and better infiltration values from the chlorides. So we are using temperatures and chlorides right now to constrain infiltration and percolation flux. We need to add strontium, we need to add sulfate, we need to add other conservative species to allow us to more pin down the percolation flux, which is very important for seepage calculations.

PARIZEK: That's different than the chemical interaction, yeah, implications such as the silica discussions you heard of.

BODVARSSON: Yes.

PARIZEK: It excludes that.

BODVARSSON: Right.

PARIZEK: You then cite natural analogs, and I don't think any were in the presentation. You mentioned examples of the kinds that you're using.

BODVARSSON: No, they're not the analogs we're using for UZ flow and transport model. Jean Younker mentioned this before. Number one priority in my view is to explain the rapid movement of radionuclides that have been observed at Hanford, INEL and NTS, because I believe you can never have confidence in our models unless we explain those. That is the emphasis right now, all the natural analog studies, in addition to the Pena Blanca.

Pena Blanca will be directly used in this UZ model. We are also planning to use geothermal analogs especially for the silica case that you mentioned, because I think we can use geothermal analogs to get reaction rates on calcites and silica and use that to bound processes, including the silica dissolution and precipitation.

PARIZEK: So those are the main ones that you see useful?

BODVARSSON: That's the main ones. Do you agree with that, Abe?

PARIZEK: The Board has received some comments from a Dr. Donald Baker, and particular a groundwater issue that was published in this July/August issue was a paper by Baker, Arnold and Scott, and there, they challenge and criticize the program for the mathematical approach that was used to model the unsaturated zone. Baker argues that the use of an arithmetic standard means for describing the block hydraulic connectivity numerical models is incorrect, and can lead to substantial errors, and recommends that the program needs to do this, otherwise maybe you're creating error upon error in the total analysis.

And I guess the Board is looking for some response to that kind of criticism. Do you feel like the Baker article is critical and is valid, or is it really a skimming problem, and as a result, you can't put in the level of detail that he implies on grid spacing it takes to perhaps deal with his concern? So do you have any comments at all on Baker's article?

BODVARSSON: Yes. Yes. We are aware of his concerns, and I don't have a personal website, but if you want to know about me, you can go to his website. I would not tell my mother the location of that website.

What Dr. Baker says, and I don't know where I can stand so you can see this, Dr. Baker did a Ph.D. thesis on rating schemes between grid blocks. And when you fix a--in two grid blocks, you can analytical belie an expression, which he did, that says this is the best expression to use to argue its permeabilities, mobilities, whatever the heck you want to argue.

The fact of the matter is that we have studied these rating schemes for ten years, and everybody studies rating schemes. They are for our problems immaterial. But we decided anyway, since the Board was concerned and Congress is going to get it, that we decided to do a case exactly like his. His work, as far as I know, as far as I've seen, only considers homogeneous porous mediums that we cannot use in our dual permeability models, but we may be able to modify it.

But the fact of the matter is we did the very extreme case of a pulse moving down through the mountain in steady state. We did steady state with the most of our results identical to his. We used his scheme, put it directly into our models, and for steady state, they are identical, totally identical. So we decided to do some--

PARIZEK: That's your grid spacing, your model, but with his scheme?

BODVARSSON: Right. Then we decided to do a pulse, because he is mostly interested in pulse, so we did the pulse of 100 millimeters in a 10 millimeter background, and the results are practically identical, too. And we have a little five page write-up that has ten pictures, all of which show that the rating schemes are immaterial for that problem.

PARIZEK: Okay. So you've considered it and it looks like it's a non-issue?

BODVARSSON: Yeah. As far as I'm concerned, it's a non-issue. I'm going to send that information to DOE, but I'm not going to put my name anywhere.

PARIZEK: If it's not publishable, maybe it's not credible.

BODVARSSON: I don't want to--you know, my feeling is whatever they say back, the reply is always going to come back.

VAN LUIK: I was going to make a different comment, but let me talk about the Dr. Baker thing. We are receiving, or are in the process of receiving an unsolicited proposal from Dr. Baker to further investigate his work, and we are going to put together a team of experts to address it. And Bo will not be part of that team, since he's already implicated on the website.

The thing that I wanted to stand up and correct is a minute ago, I think the question was do we agree that this model is valid, and I think my head kind of bobbed for some reason, and the record was said to say that I shook my head.

We don't agree that the model is valid. We agree that the activities that are underway and are planned will give us a good handle on how correct this model is for the purpose at hand.

On the other hand, the reason that we can nod our heads affirmatively at this time is that it looks like the trend is that all of the work that's being done now is going to cut back on the percolation flux that is predicted. And so we think that the model that he's doing the 30 flow fields on now is actually a conservative one compared to what it will be a couple of years down the road.

So we have pretty good confidence that this is the right way to go, but I hope that neither Bob nor I were interpreted as saying yes, this model is valid.

CRAIG: Okay. On that note, we're going to have to move on, Bo. Thank you very, very much.

BODVARSSON: The seepage.

CRAIG: Well, we have a time problem.

VAN LUIK: I'm sure Bo can do it in five minutes.

BODVARSSON: Five minutes.

CRAIG: All right, we'll give you five minutes.

The price you pay for inviting questions in the middle of your talk.

BODVARSSON: Yeah, it's my fault.

Okay, seepage calibration model, real quick. Stephan Finster at LBL just finished one of the AMRs on seepage calibration. I am very proud of his work. I think he does excellent work. He uses mainly a three dimensional heterogeneous field with different permeabilities. He uses that to match all the data. That includes memory effect, because if you have a pulse right after another pulse, it remembers the first pulse, and looking at seepage threshold, that's the main emphasis of this work, plus making a calibrated model for PA.

He used four different models, 2-D and 3-D homogeneous and heterogeneous models to compare the results. He uses a lot of statistics to match the data, and then he used another data set to validate his results. He calibrates mainly the alpha van Genuchten parameter and the fracture porosity. These are the four different models, and you see they have fairly similar fracture porosities from .1 per cent. There are a little different alpha because of the three dimensional nature. So this should be more accurate than this one or alpha for the PT models.

He just completed the results with an AMR because the computer has been cranking and cranking and cranking on a 3-D heterogeneous match that's shown here. These are the various tests, and the 2-D homogeneous, heterogeneous, and the you see they are all very, very consistent results.

Now, what does this mean? Then he uses "validation" when he takes another data set, uses the calibrated model, and in this case, I guess the predicted is the red one, the mean is this gray one here, or vice versa. And in most cases, he concludes that the predicted seepage percentage is consistent with absolute values on a 95 confidence basis.

Finally, he did Monte Carlo simulations to look at the seepage threshold, and this slide was done before the AMR was reviewed, actually Chin-Fu Tsang was sitting there, was my technical reviewer for this AMR. He concluded that the seepage threshold for the middle non-lithophysal unit or the four meter niche is 1000 millimeters per year, which I think is a major conclusion which is based on a lot of simulations, as you see here.

Now, what does that mean for the lower lithophysal? What does it mean? Of course when you have a bigger niche like 5.5 meters, this may go down some, but this is a very large value and could have huge impacts, at least I think personally.

And that's it in five minutes.

KNOPMAN: Is the AMR for that done?

BODVARSSON: Yes. The AMR, you've got a copy of the

AMR. All the Board members, I sent two AMRs.

CRAIG: A quick question from Debra?

KNOPMAN: I just want to make sure that we have copies of these viewgraphs, these new viewgraphs.

BODVARSSON: Yes. Do you want the one on Baker?

KNOPMAN: Yes.

BODVARSSON: Okay.

CRAIG: Thank you very much, Bo. And now we turn to Joe Farmer.

I see a special session this evening, or something, on the 1000 millimeter flux. Clearly, we could talk about that for a long time.

FARMER: First of all, I'd like to thank DOE, the project and the Board for the opportunity to speak. It's certainly a distinguished group of people on the Board, and of course it's a privilege for all of us to have your attention, and do appreciate the opportunity to be here.

The title of this particular presentation is the development and validation of realistic, realistic I hope, degradation mode models for the waste package and drip shield.

This is basically a cartoon of the current EDA II design. And of course in the EDA II design, we're using Alloy-22 as a corrosion resistant outer barrier. We're using 316 NG, both as a structural support, and something that hasn't been mentioned much to date, but also as a type of radiation shielding. We have Titanium Grade 7 that we're using as a drip shield over the outside of the waste package. This will protect the waste package both from rock fall as well as from dripping water.

There have been some clever but unmentioned things taken into account in the design of this particular system. I know the engineers have taken special care to isolate the Titanium Grade 7 drip shield from the carbon steel invert, and of course this is very important because if you get galvanic coupling between a carbon steel invert and the Titanium drip shield, you could get cathodic hydrogen charging, and they have in fact designed this feature out of the system. So that isn't a concern in the current design.

And, of course, if we have backfill over the drip shield, we also don't have to worry about rock bolts and netting and other things falling down on the top of the drip shield. This has been a concern that's been raised in the past, but I don't think it's a concern that we have at the present time.

Another feature in the design not mentioned yet is the fact that we're using Alloy-22 clad waste package supports, and this is a very important feature because it tends to give us an Alloy-22/Alloy-22 crevice in this particular region, and as you'll see in some of the subsequent viewgraphs, this will substantially limit the possibility for having a very bad aggressive environment in this crevice region.

This is an integrated mechanistically based degradation mode model, and in essence we're using the same general type of schematic for the Titanium Grade 7, the Alloy-22 and the 316 NG.

In this particular integrated model for the waste package outer barrier, we account for the local environment on the waste package surface. We also have a number of thresholds built into the model so that we can switch from one type of failure model to another.

We have a number of mode specific penetration rates that we sum up to give an overall penetration rate. Unlike the models that we used in TSPA/VA, we're now incorporating the ability to deal with phase instabilities in the Alloy-22, which is an important issue that I believe we're adequately addressing at this particular point.

We're accounting for various types of manufacturing defects, such as flaws that could promote stress corrosion cracking. We have two competing models for stress corrosion cracking, one that we've been using historically, and when I say historically, probably over the last two or three years, that's based on a threshold stress intensity factor.

In this particular case, it's assumed that if the stress intensity at the tip of a flaw exceeds the critical threshold stress corrosion cracking, we will in fact promote and propagate the stress corrosion crack through the wall of the container.

A competing model that comes from the nuclear industry is known as the film rupture model. In this particular case, it's assuming that even without a pre-existing flaw, you can in fact nucleate a stress corrosion crack and have that propagate at a relatively slow rate through the wall of the container by periodically rupturing a film at the crack tip. And since there is some disagreement as to which of these models is best, we're pursuing both in parallel.

Today, I'd like to discuss with you some of the general strategies that we're using in an attempt to validate our models. In most cases, the type of validation we're doing is in essence using independent measurements in an attempt to corroborate our predictions and our models.

We're also doing some bounding analyses, and looking at the results of these bounding analyses to see if they pass the Ho-Ho test, or if they are at least in a regime that makes sense to us.

The examples that we'll be covering with you today are general and localized corrosion, crevice corrosion, stress corrosion cracking, and aging and phase stability.

The first example of using corroborative data will be where I show you some of our very low general corrosion rates, and I'll show you how we've used a cutting edge technique, Atomic Force Microscopy, to confirm and validate that those corrosion rates are indeed as low as we believe them to be, and as low as we're modelling.

I'll also mention to you how we're using cyclic polarization to validate or confirm that these materials are passive and stable over very broad ranges of potential, and a variety of aggressive environments.

In terms of crevice corrosion, I'm sure the Board remembers from a few years ago we were out calculating exactly how severe the environment could be in various crevices. And the Board correctly recommended to us that maybe it would be wise to go out and actually try to measure these. So at this particular point, I'm happy to say that we've taken that advice to heart and we have gone in and made in situ measurements of crevice pH and found that our transport calculations were pretty much dead on the money.

Stress corrosion cracking models, we have two competing models, and I'll say a few words about the types of data that we're collecting both to fit the parameters in those models, and also the types of testing that we're doing to validate and show if those models are adequate for predictive purposes.

Another more important feature that I'll discuss in the stress corrosion cracking model area is the fact that we are concerned that any stress corrosion cracking may be unacceptable. So we proposed a process several months ago that we believe could perhaps completely mitigate stress corrosion cracking, perhaps even eliminate the need for stress corrosion cracking models.

I showed some preliminary data with non-waste package materials in Beatty. We now have data with Alloy-22 welds that are representative of the types of welds we're going to have in the waste package. I believe we're validating this mitigation technique as a means for perhaps eliminating stress corrosion cracking as a major concern.

Over the last 18 months, two years, we've collected a lot of data on aging and phase stability. We've also built up a theoretical capability for predicting time/temperature transformation diagrams, as well as rates of precipitation at various intermetallics. So I'll try to show you at least Anna Whitman's sampler approach, how we're trying to use the transmission electron microscope to go in and validate and confirm these phase stability models.

Before I get into discussion specific degradation modes, I of course mentioned to you in the previous chart that we've tried to account for how the local environment on the waste package surface differs from the groundwater or the near-field environment. We of course can calculate what type of evolution we have in the local environment on the waste package surface using some of the geochemical codes such as E2-36. But, again, as recommended by the Board, we've now gone in and done a large number of experiments where we actually do evaporative concentration of electrolytes.

This is just one example. But in this particular case, we've evaporatively concentrated 5000X J-13, and you can see that after we remove about 90 per cent of the water from this initial starting solution, the electrolyte evolves into a sodium potassium chloride nitrate solution with some residual carbonate buffer.

In this particular case, the boiling point is around 112 degrees Centigrade, and it has a pH of 12. We can go to even higher boiling points and more concentrated electrolytes, but we believe a 90 per cent water removal is perhaps more aggressive than a fully saturated solution, because we have still quite a lot of dissolved oxygen. Without dissolved oxygen, your corrosion rates go to a very low level. So to go to a fully saturated solution is not necessarily going to the most aggressive condition.

We also have a variant test medium based upon this 90 per cent water removal, which we refer to as SSW. In essence, it's a sodium potassium chloride nitrate solution with a boiling point of 120 degrees, much higher than this, and without any buffer present. And we believe that's probably certainly pushing the envelope in terms of how aggressive a medium could be.

We're relying a lot, at least on bounding our corrosion rates, with data from the long-term corrosion test facility. Of course, we have to make sure that the electrolytes used in the long-term test facility are saturated with oxygen. If they are not, that means that the rates we're measuring would be not as conservative as we would like them to be. We've gone in in an attempt to validate our measured dissolved oxygen and compared them to published data for synthetic geothermal brines, and based upon these comparisons and other data, we believe that we are in fact saturated in oxygen in the long-term corrosion test facility. So any data coming out of that facility should be conservative in nature.

We use weight loss and dimensional change of several hundred Alloy-22 and Titanium Grade 16 samples as a way of inferring what we believe the bounding corrosion rates are for the waste package materials.

In this particular case, we see that the corrosion rates, or general corrosion rates that are calculated from these weight loss and dimensional changes for both Titanium Grade 16 and Alloy-22 are, in essence, a Galcean (phonetic) distribution of measurement error.

Now, that sounds bad at first, but what we're really saying here is that the general corrosion rates fall below the limit due to this measurement error. And such low corrosion rates will not be life limiting.

As we mentioned before during Jean's talk and some others, in the case of Titanium Grade 16, which is an analog of the Titanium Grade 7 that we're using, we see that the general corrosion rate is never observed to be greater than around 350 nanometers per year, or .35 microns per year. And, of course, this would give us a waste package life--or I'm sorry--a drip shield life much longer than what we would need to meet regulatory requirements.

In a similar fashion, if we look at the highest observed rates for Alloy-22, which are bounded by this distribution of measurement error, if you will, we can see that the highest observed rate of 150 nanometers per year, or .15 microns per year would never limit the life of a waste package.

Now, we realize that we have some skeptics in the audience, so we didn't want to just go out and tell you that we're making measurement error measurements, so we realized early on that we had to take some steps to prove to you and show that these general corrosion rates are as low as we say that they are.

Here in the upper left-hand corner, you see a surface image, an Alloy-22 surface image with Atomic Force Microscopy. Here, you can see some of the machining marks on the surface of the Alloy-22 as it comes from the mill. We then do a vapor phase exposure of this sample in the long term corrosion test facility, and there's not exact registering between this machining mark and this one. You know, it's, on a nanoscale, it's very hard to get these things to register. But the topography is representative.

But at any rate, we do a one year exposure of this sample at 90 degrees Centigrade in a simulated acidified water, which is about 1000X J-13 at a pH of 3, and you can see the onset of oxidation and corrosion with perhaps some scale formation. But the important thing is in no case does the topography increase or exceed .3 microns per year, or about 300 nanometers per year.

So certainly the general corrosion that we image with the Atomic Force Microscope is consistent with the limits that we set with these weight loss measurements. So this is one way that we go about validating or confirming these general corrosion rates, or the limits that we are setting on general corrosion with the weight loss.

This is another sample exposed to the same medium. In this case, it is a liquid phase exposure. If you look at the portion of the surface that is below the silica scale, once again, you see that the general corrosion and oxidation that you infer from the change in topography is less than about 150 nanometers per year, or .15 microns per year.

So, again, this is confirmatory and would tend to substantiate our claims that the corrosion rates are in fact quite low. We see these glacial type deposits form on the surface of these Alloy-22 samples when we put them below the water line, and we use low angle x-ray defraction with a Regatu (phonetic) stage to show that these deposits are basically silica. And I think this gets back to one of the person's comments having to do with immobilization of silica.

So we've actually been toying with the idea that maybe what we really have here is a silica coated waste package that extremely corrosion resistant. So this is probably working to our advantage.

Now, of course, the reason that Alloy-22 and Titanium Grade 7 is so corrosion resistant is because these materials exhibit passivity over tremendously broad range of electrochemical potential. As we do cyclic polarization or potentiodynamic measurements, we go from the corrosion potential up to a higher or more anodic potentials where we might start expecting the breakdown of either water or the passive film on the material. We see that the separation between the corrosion potential and the threshold, or possible threshold potential, is very large, 1000, 1200 millivolts.

This tremendously large separation between these two defining potentials is a quantitative measure of exactly how corrosion resistant this particular material is. There's no plausible way that I can think of to ever get up and do this regime where you might start arguing that you have some type of breakdown of the TI 02 passive film.

So certainly Titanium Grade 7, Titanium Grade 16 are very stable in these environments where we're testing. In this particular case, it's a test in the simulated saturated water, saturated sodium potassium chloride nitrate solution at 120 degrees Centigrade.

We do similar measures with Alloy-22. In this particular case, the SSW at 120 degrees Centigrade. Here again, you see that you have a very broad range, or a very broad potential separation between the corrosion potential and the threshold potential. And, in fact, this threshold potential is the onset of oxygen evolution. It doesn't really define the catastrophic breakdown of the passive film. But because of the nature of the measurement, we simply know that if the passive film does break down, it's somewhere above this level.

So you can see that we have passivity over an extremely broad range of potential, and the only way we can destabilize this passive film is to somehow magically push the corrosion potential up to that level where we will break down, thermodynamically break down the passive film.

This type of behavior that you see to the Titanium and the Alloy-22 is in very sharp contrast to what you see for other materials, such as 316L. And 316L, for all practical purposes, is about the same material as 316 nuclear grade, 316 NG, which is the material that we're going to use for the structural support.

In this particular case, you see that you can have a catastrophic breakdown of the passive film at potentials relatively close to the corrosion potential, and there are plausible mechanisms for pushing the open circuit corrosion potential from this level, up into regimes where you would get this catastrophic breakdown of the passive film. And this, of course, is the reason that the engineering on the project decided to use these types of materials for the drip shield and the waste package outer barrier, and not the 316.

But even though we're not using this particular material for its corrosion resistant properties on this 10,000 year time frame, it is in fact quite a good structural material.

There are some unusual effects that we've observed in Alloy-22 and we feel like it's our professional and ethical responsibility to point all of these warts and bumps out to you, and this is basically what we're doing here. We test Alloy-22 in a simulated concentrated water. Again, this is about 1000X J-13. We still see in this particular case that we have to push the potential up well over 700 millivolts to get a breakdown or failure of the passive film, if you will.

However, there is a redox couple that is due to some oxidation state in the passive film. In a perfect world, you'd probably prefer not to see this redox reaction, even though it doesn't seem to do anything in terms of de-stabilizing the passive film. But as I'll show you in a second, we still believe that there's no plausible way of getting up into this redox regime.

And, of course, we've confirmed that this is a redox couple in the oxide film and not in the electrolyte by comparing an electrochemical scan for a platinum standard. You see the peak on the Alloy-22, but not on the platinum in the same electrolyte.

If we set at the potential that coincides with the onset of this anodic oxidation peak, we basically see that we have an electrochemical reaction where we're probably changing the oxidation state in that passive film, but eventually we get conversation of the passive film, and the current density that we measure returns to around 4 microamps per square centimeter, which is representative of a typical passive current density that we observe with Alloy-22.

So this basically is evidence that even though there is some type of redox reaction here, that the passive film is intact and stable.

So we have two types of thresholds that we can define with Alloy-22, one due to the catastrophic breakdown of the passive film. This is a region that we absolutely want to avoid because if we go above this level, you remove the protective oxide film and you can get dissolution of the metal. And then this other, I would call sort of a nuisance peak where we might get some sort of temporary redox occur. And to be conservative, we're actually using this redox peak in the case of the SCW electrolyte as defining the maximum potential that we're willing to accept. And then, of course, we also go out and measure corrosion potentials.

Now, I mentioned to you that we're basing a lot of our model on these corrosion and threshold potentials. We have to assure that we don't have some magical means of pushing our open circuit corrosion potential of any of the waste package materials into regimes where we expect harm to come to the waste package.

One technique, or one way that we might push the open circuit corrosion potential into a region of trouble would be from gamma radiolysis. Gamma radiolysis generates a number of species, but the one that primarily affects the electrochemical potential is hydrogen peroxide. So we go in and actually investigate the effect of hydrogen peroxide on the open circuit corrosion potential.

A number of years ago, some of you may remember this, at Livermore, we actually used a cobalt 60 source and gamma pit studies to go in and quantify exactly how much impact the gamma field had on the open circuit corrosion potential. Since we don't have the time or the resources in our current environment to go in and repeat the gamma pit studies, we have instead mimicked the effects of gamma radiolysis using hydrogen peroxide additions.

Based upon these measurements, we believe that we're going to be able to screen out the gamma radiolysis as a serious threat.

Here are some experiments where we have looked at the change in the open circuit corrosion potential as a function of hydrogen peroxide addition. The numbers above the curve represent steps in hydrogen peroxide concentration in parts per million. So here we have zero, 8, 16, 24, 32, up to 72 parts per million hydrogen peroxide in the electrolyte. And, of course, we basically titrate this over some period of time, and we simultaneously monitor the open circuit corrosion potential.

In the case of the simulated concentrated well water, J-13, we see that the maximum corrosion potential that we ever achieve by these hydrogen peroxide additions is less than zero millivolts versus the silver silver chloride reference electrode.

In the case of that anodic oxidation peak I showed you, you would have to have another 200 millivolts of potential before you could even get a redox change in the film. You'd probably have to have another 700 millivolts above this maximum change in corrosion potential before you could get into a regime where you would have localized breakdown of the passive film.

So through experiments like this, we believe that we can more or less bound the effects of gamma radiolysis, and hopefully use that as a means of taking that off the table in terms of being a major concern.

We, of course, perform these experiments on all of our various test media. Here, we have a similar experiment performed with simulated acidified water, and in this particular case, we see that the maximum anodic potential that we can achieve is 150 millivolts. Again, in this particular case, in order to destabilize the passive film, we would have to be well above 700 millivolts. So we have probably well over a 500 millivolt margin, and I don't think there's any plausible way of getting there.

So this data goes to make the point that Alloy-22 is a very stable material indeed.

We've spent a lot of time over the last few years worrying about crevice corrosion, and the TSPA/VA design when we had the carbon steel outer barrier, this was quite a serious issue because as we would tend to corrode through the carbon steel barrier, we knew that we would form a crevice between what was left of the carbon steel and the Alloy-22 surface, and that ferric chloride solution, which would be quite acidic, could be harmful to the Alloy-22.

In the current design, we know that we're still going to have crevices that are going to form in these mineral deposits, corrosion products, and even between the outer barrier and the inner barrier if you have some breach of the outer barrier. Also between the waste package and supports.

In a crevice, as most of you realize by now, we can have a very low pH, because the dissolved metal in these occluded geometries can hydrolyze to give you hydrogen cations, and the field-driven electromigration of chloride into these regions will tend to further exacerbate that environment.

This crevice environment can accelerate general corrosion, pitting, and stress corrosion cracking. Now, of course, the successful defense of the waste package requires that we develop a thorough understanding of that.

As we showed you in Beatty, we've now gone in and actually physically measured the crevice pH in these environments, and of course this was the recommendation made to us by the Board.

What you see in the upper left-hand corner is that in the case of 316L and 316 NG, at relatively low polarizations, low electrochemical potentials at the mouth of the crevice, we can achieve almost spontaneous low pHs. So if we were going to form a crevice with 316 in the waste package design, it could be quite threatening.

However, if we go to Alloy-22, which remains passive over a very broad range of potential, up to around 1000 millivolts, we see that the pH is not nearly as severe. For example, at around 400 millivolts, the pH never drops below 6. So in these passive crevices formed from Alloy-22, we do not believe that the crevice environment is going to be as bad as it would be with material such as 316 NG.

In the lower right-hand corner, you see the crevice current that corresponds to the measured pH. In this particular case, we see that we have to go to around 1000 millivolts before we get catastrophic breakdown of the passive film inside the crevice. And at that particular point, we see a large increase in the current going out of the mouth of that crevice.

In this particular picture, you see a special electrochemical cell that we have built and operated to go in and make these particular types of pH measurements. This particular slide shows you two samples used in this artificial crevice. The one on the left was polarized for several weeks at 400 millivolts, and of course you see virtually no attack of the Alloy-22. The one on the right was polarized at 1100 millivolts at the crevice mouth, and in this particular case, you see both a lot of oxidation of the Alloy-22 surface, and a lot of severe crevice attack along the leading edge of a mass that was used to define the front end of that crevice.

And as we look at this creviced environment up close, again we see virtually no noticeable attack of the Alloy-22 at 400 millivolts. But at 1100 millivolts, we see that the crevice attack can be severe indeed. So the lesson learned of course is that you don't want to push these materials above their critical or threshold potentials. And that's why a lot of the current model is based on these types of thresholds. They're incorporated into the TSPA/VA model at this particular point.

As Jean mentioned yesterday, it's important that we use corroborative data. So in addition to doing calculations first of all, based upon transport, and calculating what these pH levels should be, we use in situ sensors to measure the pH, and then we go out and use other techniques, such as inserting indicators papers into these crevices.

In this particular case, you can see that under open circuit conditions, we have a neutral solution in this particular crevice. But as we polarize it at 800 millivolts, it starts to acidify, and of course the paper turns a corresponding color, a color that would correspond to a pH of somewhere between 1 and 3.

And just to show you other corroborative data, we performed similar experiments with 304 stainless steel, and in this particular case, once we polarized the mouth of the crevice, you not only see a general acidification and a passive crevice, you start seeing the nucleation of pits and the acid oozing or flowing out of the mouth of those pits. Of course, this is again the reason we didn't pick a 300 series stainless steel as the outer barrier of the waste package. But we are in fact doing a lot of corroborative measurements like this to validate our models and make sure that our concepts are correct.

And this, of course, is an old model prediction that I think I showed you a couple of years ago, and I think the bottom line here is that we're now measuring at 800 millivolts a pH between 2 and 3, and these were our model predictions at that particular point in time. So I think the data is bearing out that some of our earlier concepts were in fact correct.

To summarize, we look at the crevice corrosion of the Alloy-22. We have two boundaries that we worked between. If we have buffer in the electrolyte that makes up the crevice solution, we get little or no suppression of the pH in the crevice. If we remove that buffer and work, let's say, with an essentially saturated chloride environment, we can get pH suppression in the crevice, and at the point where we get a complete breakdown of the passive film, the pH can go to a very low level.

But at reasonable polarizations, let's say 200 to 400 millivolts, the amount of pH suppression we get in this crevice is not great. If, in turn, we have a 316 crevice, we can get to much lower pHs.

One of the reasons that we worry about pH suppression in crevices with Titanium is that the low pH, the high concentration of hydrogen ions, coupled with a cathodic polarization, can in fact drive hydrogen into a crevice region.

In this particular case, we see hydrogen profiles determined with secondary ion mass spec in a Titanium Grade 16 crevice. These are ratios of counts per second for hydrogen and Titanium. I haven't converted these to parts per million. But the bottom line here is that we can use SIMS as a method of determining the maximum hydrogen absorption in these Titanium based crevices.

What we've observed, once we use calibrated signals, is that the absorbed hydrogen remains below around 1000 parts per million. In order for us to get hydrogen induced cracking, even in a Titanium crevice, we have to be above the threshold of 1000 parts per million hydrogen.

So this is the type of data that we're using to go in and determine both parameters in the hydrogen induced cracking model, and also set thresholds and to some extent validate models and concepts.

CRAIG: Joe, you've now used your full allotted half hour.

FARMER: Can I sit down now?

CRAIG: No, no, we're not in a crisis mode yet, but we want to get back on schedule.

FARMER: Okay. Sure.

BULLEN: Mr. Chairman, I would suggest we take time from the panel and finish the presentation.

CRAIG: Well, I'm not proposing to stop the presentation.

BULLEN: I mean, if we have to run over with Joe, I would just suggest we take time from the panel, maybe 10 or 15 minutes.

CRAIG: Okay. Why don't we push on and see where we are.

BULLEN: Okay, that's fine.

FARMER: All right. Well, let me I guess just to basically put back up my road map, and I apologize for the somewhat chaotic nature of the presentation, but I believe I at least have given you some flavor of the types of work that we're doing to go in and look at the local environment on the waste package surface. I've shown you some of the data that we're using to determine these mode specific penetration rates. We of course are going in and physically measuring these corrosion and threshold potentials as well as experimentally and numerically determining these minimum possible pH levels that can form in crevices.

So we're trying to basically go in and measure all the pieces of this puzzle. The things that I haven't shown you yet are over on the right-hand chart, right-hand side of the chart. We're doing a lot of work to go in and look at the phase stability of Alloy-22. This is a very important issue. And we're also doing a lot of work to shore up these stress corrosion cracking models.

This is something that we didn't account for in TSPA/VA, and it turns out in the current waste package design, this is probably going to be one of the most serious concerns that we have to worry about.

So now before I sit down, I'd like to just say a few words about the phase stability and the stress corrosion cracking and how we're going to mitigate that.

We actually, as I said before, we have two competing stress corrosion cracking models, one based on a threshold stress intensity factor, and another based on the film rupture model. To both validate and also determine some of the parameters, we're using the double cantilever beam method. This particular method has been illustrated for you before.

We've now placed a contract to General Electric Corporation. We're using the reverse DC method of Pater Andresen to determine the crack propagation rates as a function of stress intensity and various environmental parameters. So we are, in fact, looking at two alternative models to address the stress corrosion cracking issue.

We have done a stress analysis of the unperturbed waste package. We've accounted for three basic sources of stress, one due to mass loading of the container, another due to the shrink fitting or thermally enhanced fit process, and finally, we've looked at the stresses due to unannealed weld stress.

As you know in the waste package, after you load the fuel in, you can't heat the waste package above 350 degrees Centigrade because of the limits on the cladding of the fuel. So we can't use a thermal process for annealing out the weld stress. We have to come up with some other technique for doing this if we want to mitigate the driver for stress corrosion cracking.

At Beatty, we mentioned to you that we were looking at laser peening as a method for mitigating these residual weld stresses that are the driver for stress corrosion cracking. We had some preliminary data with a 4340 steel, and had actually looked at using double pass laser peening as a method of driving compressive stress deep into the waste package weld. And, of course, if you can introduce compressive stress, it counters the tensile stress that would tend to drive the stress corrosion cracking.

These are some data for prototypical waste package welds. These measurements were made .2 inches from the fusion line. This is made right on the centerline. Here, you can see in this particular invention, positive stresses are tensile negative, or compressive.

So, in essence, you see that in the un-peened waste package weld, we had relatively high tensile stresses. In this particular case, the yield stress is around 55 ksi. After doing laser peening, we can push those tensile stresses down into the compressive region. And, of course, if we convert the stresses in that waste package weld from tensile to compressive, we can in essence mitigate stress corrosion cracking and prevent it from occurring. So it's sort of like inoculating someone to make sure they don't get the chicken pox perhaps.

A similar case over here right on the centerline. You start out with relatively tensile stresses, but after doing laser peening, we basically can drive those into compression. And I can tell you a little bit about the laser and the system if you want to ask during questioning.

We have theoretical models to now deal with the phase stability and the precipitation kinetics in Alloy-22 and other materials of interest. The two codes that are being used are THERMO-CALC and DICTRA. These are a phenomenological codes that can predict energetics, regions of stability and metastability, as well as phase transformation rates limited either by kinetics or diffusive transport.

And, of course, in some of these models, you lack some of the thermodynamic data that you need, so we're using an electronic structure based approach to augment the database so that we can do the jobs that we need to do.

As you've seen before, we can in fact precipitate intermetallic particles. These are generally Ni2, CR Ni2 MO type particles. These intermetallics are bad because they can deplete alloy elements that are responsible for the passivity of Alloy-22 and open up areas for localized attack of the materials. These precipitates can also embrittle the material and make it more prone to failure if there's a rock fall. So it's very important that we understand the precipitation kinetics.

We're actually going in and using the volume fracture of precipitate as a function of time and temperature to validate our kinetic models.

Here, you can see a material that's been purposefully aged to 1000 hours at a relatively high temperature. And if you age these at a long enough time and a high enough temperature, you can eventually completely cover the grain boundaries with intermetallic precipitates.

We have started to collect enough data so that we can in fact construct empirical time/temperature transformation diagrams. We're using DICTRA to go back in and do a more precise job of defining these boundaries between regions of partial grain boundary coverage, complete grain boundary coverage, and also to define regions of long-range ordering.

The bottom line here is we're going to be operating our waste package somewhere below 350 degrees Centigrade, so in our particular case, we don't believe that phase instabilities in the material will be a life limiting problem.

We've also gone in and started to do kinetic measurements. These lines represent the point when you would first initiate grain boundary precipitation, and this other line represents, for example, when you start having precipitates form in the bulk material. The red line represents the point when you've completely covered the grain boundaries with precipitates.

So we are both experimentally and theoretically looking at the precipitation kinetics in these alloys to prove that they have the stability that we need.

In summary, we believe that validation is an essential part of model development and requires quite a lot of time to discuss in a presentation like this. I've tried to give you four examples of model validation, one related to general and localized corrosion, another having to do with crevice corrosion, some having to do with stress corrosion cracking, and finally, some having to do with phase stability.

Some preliminary conclusions. At the present time, we don't believe that the waste package is going to be limited by general corrosion. We don't think that localized corrosion is going to be a significant problem with this particular material. Preliminary data indicates that phase stability will be acceptable.

We are, of course, as I mentioned, focusing on mitigation of stress corrosion cracking at the final closure weld. We have two competing models for stress corrosion cracking, and we're doing a lot of work with the laser peening as a way of eliminating the tensile stresses that would tend to drive that particular mode of failure.

We have a new design. Two materials were brought on board with the new design, Titanium and 316. Tests on these materials for all practical purposes have just begun. We've been testing probably less than six months with these materials, and need a lot more data.

We know that we have at least two fabrication processes that are going to require some additional research and development. We have a thermally enhanced fit of the Alloy-22 over the 316 NG, and we need to understand very well exactly what type of tensile stresses will be introduced into the Alloy-22 as a result of that thermally enhanced fitting process. And we also realize at this particular point that it's going to be important to bring on board some of the state of the art techniques, such as laser peening, to mitigate stress corrosion cracking.

And I would like to point out that the peening is not a toy box type process. It's actually being used to treat turbine blades on some very high performance aircraft that are very important to us, and it's also being used to do peening on some gears that have equal importance. So it isn't just a sandbox process, and it's been commercialized.

So I'll be happy to answer any questions.

CRAIG: Okay, wonderful. We have time for some discussion. Dan Bullen?

BULLEN: Bullen, Board. Actually, Joe, I want to compliment you to begin with, because it's always very nice for people to acknowledge that we've made suggestions and that the DOE and the M&O contractors have gone out and actually done the things that we might think would be important, and then to have those results come back to us and say, well, this is what you told us you wanted to do, and we did it, is always a little bit reassuring.

Now, unfortunately, that never comes free, and so I know it costs money, and you probably had to do things that otherwise you might have done because of that.

I have a number of issues that I want to talk about. I guess the first one will always be radiolysis. And as I go back to the radiolysis issues that were raised on Figure 9, we started talking about the polarization curves.

FARMER: Okay.

BULLEN: The question that I have for you deals with the fact that if you add the hydrogen peroxide--actually I guess it would be subsequent to that. It was a little bit farther down. Your Figure 12, where the radiolysis--as you titrated in the hydrogen peroxide.

FARMER: Right.

BULLEN: The question that I have for you is in an aqueous environment, this all makes sense. But in a thin film environment underneath the drip shield, if you're trying to take a look at the condensate that's there, and as you introduce, you also have hydrogen peroxide that would be there, which is the detriment, in the radiolysis environment, you're going to have other actors that will be there.

Now, for the Titanium, the nitrates and the nitric acid probably are who cares, because that's actually a beneficial breakdown, but are there any other things that might jump up and bite you? Are there any surprises you'd expect to see? And if so, are there tests that you think you could do or should have done, or maybe would want to do? I mean, before the 50 years of emplacement, you've got a lot of time to figure out how am I going to test this drip shield. And so maybe you could give me an indication of what you'd expect to try and do with respect to radiolysis testing at some point in time.

FARMER: Okay. Well, first of all, I'm putting this up not because--well, it's pretty for one thing--but the other reason I'm putting it up is because I think this illustrates the strength of the Atomic Force Microscope and why we've been using it so much.

First of all, these waste package materials for all practical purposes don't corrode. We beat on them, we dip them in lots of horrible things, and you pull them out and they basically look pretty much like when you put them in.

So if you don't have something like an Atomic Force Microscope to look at the surface, you on first appearance have a null experiment.

Now, this is a particular case where we actually observed spontaneous pitting on a 300 series stainless steel, and I unfortunately didn't have time to make a viewgraph of it, but we have similar experiments we've done where we have taken--I didn't discuss it at the microphone--but we have done some experiments where we have submersed these with hydrogen peroxide, not making potential measurements, but actually looking at the evolution of the morphology of the passive film as we dope these or add hydrogen peroxide to the electrolyte.

And, frankly, in those cases, you know, here you see a very terrible thing happening to the passive film on this 300 series stainless steel. We see nothing like this happening with the Alloy-22.

You know, Peter Bedrossian, who's a physicist who runs the microscope, will come in after he's had too much coffee and try to convince me that he's seen some change. But, you know, ten cups of the very best Starbuck's and I still can't see it.

So I think that the passive film on the Alloy-22 is quite stable, even in a thin film environment.

BULLEN: How about have you done the same for the Titanium?

FARMER: Again, this is not directly relevant, but I've shown you a lot of pictures where nothing happens, so I don't want you to get the impression that the Atomic Force

Microscope can't see anything. This is a case where we purposely took Titanium Grade 12, which incidentally is not the Titanium grade we're using, and we charged the dickens out of it at about minus 1.45 volts, and we've used SIMS here to depth profile the hydrogen into the Titanium surface, and we've looked at the evolution of the Titanium surface as we hydrogen charge it, and I show you this not because this is what our waste package is. Our waste package isn't going to look like this. But the point is if we had a problem like this, we'd sure as heck be able to see it.

You know, this is very interesting. You're actually seeing here the formation of sort of nano-hydrogen bubbles sub-surface. And the more incredible thing about this is that in this particular environment when we do this cathodic charging, when we keep the electrochemical potential on the surface, the surface remains flat. You don't form those bubbles until you release the electrochemical potential, and you start forming gaseous hydrogen inside.

So we do have the ability to see these types of phenomenon. We look at hydrogen peroxide effects on Titanium. We look at them on steel. We look at them on Alloy-22. And, frankly, it doesn't do very much at all on either Titanium or Alloy-22. In both cases, the material remains passive, and fairly boring to look at.

BULLEN: Let me change gears just for a second, and I won't take too much more time, Mr. Chairman.

On Slide 17, you say--you just glossed over it--but microbes may pose a unique threat, and I didn't see in your slide Number 3, which you actually had to put up there on the other side, anything that said MIC. Are you just grouping MIC with localized corrosion in that case? Or how do you model MIC, I guess, is the question? Where's the switch?

FARMER: Okay. Well, at the present time, we have done a lot of MIC work. JoAnn Horn, as most of you know, has headed up a very nice MIC effort in our laboratory. We have seen some very interesting biofilms form on these samples. After you remove the biofilm and start looking at the passive film underneath, again, these are very flat boring surfaces to look at.

So my gut feel from looking at them, I know there was a press conference somewhere, I can't remember exactly where it was, but it made it in the Las Vegas Sun, I think, having to do with the bugs that ate Yucca Mountain, or something to that effect. But I looked at those samples myself, and I think the holes that were seen were actually holes in the biofilm.

So we've now gone in and looked beneath the biofilm, again with the AFM, SEM, other techniques, and those surfaces do not, at least to me and others, look appreciably attacked.

Now, the thing that we are worrying about is we do have sulfate reducing bacteria at Yucca Mountain. This sulfate reducing bacteria can form sulfide. One of the key contaminants in a medium that can cause stress corrosion cracking in these nickel based alloys is sulfide. So we've pretty well I think, or we've gone pretty far down the road I think towards dismissing the hydrogen peroxide issue as a major killer, or something that, you know, the boogie man is really going to get us.

But we still have to do some work here with sulfide and sulfate reducing bacteria. We haven't quantified this yet, but we're working on it. It isn't going to be in the early revisions of the AMR, but it will ultimately be incorporated. So I guess that's the best way I can do it.

BULLEN: I'm sorry. One final question?

CRAIG: Hold on, Dan. We've got to turn--we're running out of time, and Roger Newman is a consultant.

NEWMAN: I'm Roger Newman. I guess I'm a consultant for today's purposes.

CRAIG: From the University of Manchester, and he's on the panel this afternoon.

FARMER: He knows more about stress corrosion cracking, or he's probably forgotten more about stress corrosion cracking than we will ever know.

NEWMAN: I'm actually not going to talk about stress corrosion cracking, although I think that's an interesting issue.

I wanted to just address a few things that I thought at least at first sight seem to be sort of non-conservative aspects of your testing. I just wondered if possibly you could reassure me that you've actually done the conservative versions of those.

FARMER: All right.

NEWMAN: The first one really was that your corrosion test didn't appear to be done on material containing a weld. Is that because you don't think there's a difference?

FARMER: No, actually that's a misconception, because in our long-term corrosion test facility, we have 18,000 samples. Several hundred of those samples are Alloy-22 and Titanium. I have some pictures in my briefcase I can show you of the facility. But those are both welded and un-welded samples.

In terms of our aging, we're looking both, our aging studies, we're looking both at welded and un-welded samples. Our initial cyclic polarization studies, we had to go back and do a lot of work with the base metal to kind of get the baseline data. We're now both welding samples and aging samples and comparing the cyclic polarization data we get for aged samples to that of un-aged samples.

And, of course, in some cases, you can actually see quite a large difference as you age a sample, because you form these precipitates on the grain boundaries, you can see a lot of localized attack.

NEWMAN: I mean, people that make these materials recognize that this alloy has a critical temperature for pitting corrosion, or crevice corrosion, which is close to, if not above, 100 degrees C. So it's not very surprising that you can't corrode it. However, the welded material is always assigned a significantly lower critical temperature, which can be, I believe, as low as 70 or 80 degrees. Of course, that's presumably during that testing in a very aggressive environment. But it was really just a comment about that.

Actually, I just wanted to go through a small list here. You've more or less reassured me on that one.

FARMER: Okay.

NEWMAN: The second one was that all these environments contain an awful lot of nitrate, and nitrate is a very strong inhibitor of localized corrosion of nickel alloys and stainless steel. How sure are you that there is going to be that much nitrate? Because it seems to me that your environments are sort of on the edge of a cliff between corrosivity and non-corrosivity.

You could see that actually in your results of the 316L stainless steel, where it started to pit, and then as you made the potential more positive, the pits died. And that's a classic result from, for example, Lackey and Ulig, 1966, or something.

FARMER: Right.

NEWMAN: That when you have nitrate present, the corrosion tends to occur over a range of electrode potentials. It doesn't occur at high potentials. It doesn't occur at low potentials. And so just a slight concern there that you--

FARMER: Well, what we did, we have conducted all the cyclic polarization data, and you've seen all the stress corrosion cracking data. The early tests were actually done in like 5 per cent sodium chloride at different pH levels, with no nitrate present. So we did a lot of testing in those environments. In fact, we have about five years worth of data, cyclic polarization, stress corrosion cracking data, in these sort of binary electrolytes.

What we of course were encouraged to do by this Board and others is to test in relevant environments. So one of the first things we did is to go back and take our standardized test media, which are the SAW, SDW, SCW, so on and so forth, repeat the cyclic polarization studies in those relevant test media that are based on the J-13 water chemistry, also use those test environments to repeat stress corrosion cracking measurements, and to expand those standard test media to include other bounding conditions.

Actually, it was Peter Andresen who pushed us towards these saturated environments where we evaporatively concentrate the electrolytes down to the point where we do have these sodium potassium chloride nitrate type environments.

NEWMAN: But could you have concentrated out the chloride and the nitrate together? It stays equally inhibiting as you concentrate it.

FARMER: Well, that in fact we do those experimentally. We didn't, you know, a priori, say we want to somehow run this experiment so that--

NEWMAN: I understand it's a real thing to try to simulate.

FARMER: Of course, the sulfate and the fluoride precipitate out, and eventually you can disproportionate the carbonate. So we didn't intentionally, you know, design that electrolyte. It's just sort of what we were given.

So I think that was an attempt to try to test the materials in relevant environments. And because of both the time frame that we have, you know, we're on a fairly fast track process in terms of, you know, we turn the design around and have--we had I think one or two materials before, now we have three, and two of those were on the test program. So, you know, we're trying--you kind of turn the program around on a dime, and I think we've actually done that.

But in turning the program around on a time, we have pretty well had to go through all the comments that have been made to us by a large number of review boards and panels, and we've had to pick those comments that seem to be most relevant and most dead on target, and I think to the credit of this Board, I think a lot of those comments have probably come from Alberto and Dan and Paul and others.

But we've tried to take a lot of those comments and target them very specifically, and a lot of those comments over the last few years have dealt with the relevance of the test environment. We've pushed away from testing in pure sodium chloride solutions at varying pH. So they've really pushed us towards making sure that all the tests media are directly tied to the J-13 water composition, and that there's some plausible way to get to that composition, such as evaporation.

Actually, I didn't dwell a lot on it, but you'll notice that some of the switches that we used to switch between dry oxidation, humid air corrosion and aqueous phase corrosion are actually Delaquescence points. There is a whole body of experimental data I couldn't discuss with you that's being collected by Greg Gdowski, where he actually puts very carefully and reproducibly puts salt deposits on waste package surfaces to measure these Delaquescence points so we know exactly at what threshold relative humidity we can have the existence of a truly aqueous phase.

NEWMAN: Just one more quick one, if I may.

Why did you do the crevice corrosion tests at room temperature? What was the point of that?

FARMER: Well, the reason I did them at room temperature initially is because that of course is the easiest experiment to do. And our sensors work very well. We run experiments at temperatures as high as 85 degrees Centigrade. I have sensors that I was promised would work to 127 degrees Centigrade. I'm sure they will, given enough patience and time, but the experiments of course get more difficult as you go up in temperature. We have plans to do those experiments, but we have budgetary and time limitations. So we haven't done them.

NEWMAN: And finally then, just the final thing is I don't understand why you define the corrosion potential as something that's measured over such a short period of time, because it's I think experimentally observed that the corrosion potential goes up more or less with the log of time. It's a logarithmic type of increase.

FARMER: Well, it doesn't increase indefinitely of course. There's limits to where it can go.

NEWMAN: Well, thermodynamically, it can go as high as the oxygen electrode, but I don't think it would ever do that.

FARMER: Yeah.

NEWMAN: But what concerns me, and I think this is not in any way a criticism of what you're doing, but it's more like perhaps an extension of the usual corrosion scientist's task of trying to predict the most horrible thing that can happen, is that especially if you have a bit of peroxide around, that potential you said is 200 millivolts below that critical potential where you get this transpassivity phenomenon, this molybdenum dissolution.

FARMER: Right.

NEWMAN: How do you know it's not going to get up there in a few years?

FARMER: Well, we haven't--most of the hydrogen peroxide measurements we've made to this point have been of the type that I showed you.

NEWMAN: Well, even without the hydrogen peroxide?

FARMER: Right. But we have made other open circuit corrosion potential measurements where we've monitored the corrosion potential for several months. And in those particular cases, you know, you'll see some very low frequency or very long wave lengths, if you will, change or fluctuation in the corrosion potential, but it generally doesn't fluctuate more than perhaps plus or minus 100 millivolts from its starting point. We have some data like that that I can share with you if you'd like to see it.

NEWMAN: It's funny, though, the only two real serious corrosion problems that have happened with either of these two materials in the last ten years, that's the nickel based alloys and the Titanium, were both caused by hydrogen peroxide and were both uniform type corrosion. These were discovered mainly in bleach plants and in companies that make things like toilet cleaner where they're switching to hydrogen peroxide.

FARMER: That might be a good second career.

NEWMAN: That's right. And I know that you don't have very much hydrogen peroxide, and so on and so on, but it is sort of a strange coincidence that these materials are both highly sensitive to hydrogen peroxide.

In the aerospace industry, they actually dip Titanium in hydrogen peroxide to clean it, to etch it, before they glue aircraft components together, and so on. And so there is this sensitivity. I guess I'd like to be reassured even a little bit more about how low the risk really is from the hydrogen peroxide.

CRAIG: At this point, we're going to have to take a break. I would encourage you all to come back in five minutes. Let me ask the Board to please pick up your material. Please pick up your material, Board members, because the tables have to be rearranged for the panel.

(Whereupon, a brief recess was taken.)

SAGÜÉS: We're ready now for the roundtable discussion. This is the roundtable discussion on model validation. My name is Alberto Sagüés, with the Nuclear Waste Technical Review Board. And what we are going to do first is we're going to allow the roundtable panel members to introduce themselves.

Before that, let me tell you that there are a couple of changes. Norm Christensen, who was going to be the Chair for the roundtable unfortunately had to do down to North Carolina to let the fish out, I'm told, out of an aquarium, or something like that. And as a result, I am Chairing this roundtable. And instead of Norm Christensen, Dr. Richard Parizek will take his place.

Also, another change, as it was announced earlier today, Steve Frishman is going to be replaced by Linda Lehman.

So we're going to go ahead with the self-presentations actually of the panel members, and if you could please state your name, position and affiliation, and area of expertise briefly, that will be better than my trying to do it. So we're going to start here to my right. Please go ahead.

NEWMAN: Well, you've just heard too much of me a minute ago. I'm Roger Newman. I'm from UMIST, which is a university in Manchester, United Kingdom, where I'm professor of corrosion and protection. And for these purposes, I'm a consultant to the Board. I've spent, or wasted, depending on your point of view, the last 15 years working on passivity and localized corrosion of stainless steel, and nickel alloys are more or less the same thing.

ORESKES: I'm Naomi Oreskes. I'm an associate professor in the Department of History and the Program and Science Studies at the University of California, San Diego. My specialty is the question of the stabilization of scientific knowledge, how scientific communities answer the question that's been posed many times today, which is how much information is enough. And I look at that both historically and philosophically to try to understand how scientific communities have grappled with that question in the past, and also how we might grapple with it today.

KONIKOW: I am Leonard Konikow. I'm with the U. S. Geological Survey in Reston, Virginia. I've been with them about 27 years now, and I've been working on the development and application of solutransport models and groundwater flow models primarily to groundwater contamination problems.

RUNNELLS: I suppose I should introduce myself. I'm Don Runnells, member of the Board. I'm a geochemist, retired from the University of Colorado, soon to retire from an engineering consulting firm, quite a few years dealing with the geochemistry of metals and uranium, radionuclides.

TSANG: I'm Chin-Fu Tsang from the Lawrence Berkeley National Lab. I'm the head of the Department of Hydrogeology in the Sciences Division. My main research has been heterogeneous modelling and also validation sometimes. And I was involved with INTRAVAL, DECOVALEX, that kind of thing.

APPLEGATE: I'm Dave Applegate. I'm Director of Government Affairs at the American Geological Institute. I'm a scientist by training, but a policy wonk by profession, and as a policy wonk, I can't tell you what my expertise is. There's no such thing. My experience was first spending five years in the Death Valley region studying geology there, but then spending a year on Capitol Hill working as a scientist for the Senate Committee on Energy and Natural Resources, which had a passing interest in the subject, and following it from afar since then.

LEHMAN: I'm Linda Lehman, consultant to the State of Nevada. I'm a hydrogeologist and have been involved in Yucca Mountain project and before that, BWIPP for the Nuclear Regulatory Commission in the Performance Assessment Section, and I've been doing hydrologic modelling of the saturated and unsaturated zone for the State of Nevada for about the past 17 years.

PARIZEK: I'm Richard Parizek, a Board member interested in hydrogeology, environmental geology. I'm at Penn State

University. I've been there it seems like as long as--half the buildings have been added since I came. I know too much about the sub-aspects of it, but we are still very active and supervise graduate research, and as a result, have gotten involved in the modelling of a variety of types of problems. I worked with WIPP for seven years, KBS systems panel of Tom Bickford, and then also in KBS review in the Swedish granite problem with the Board now just practically three years.

EISENBERG: I'm Norman Eisenberg from the Nuclear Regulatory Commission. I've had about 20 years experience in performance assessment at the NRC, and at DOE.

ANDREWS: I'm Bob Andrews with the M&O, manage performance assessment there, but my training is actually in hydrogeology.

SAGÜÉS: Well, thank you very much. And again, I'm Alberto Sagüés. I'm professor at the University of South Florida. My main area of interest is in corrosion of materials, and I have been also with the Board for almost three years now.

I see that in the audience we still have Bo Bodvarsson and Joe Farmer. I don't know for how long that Bo is going to be around.

ANDREWS: As long as we need him.

SAGÜÉS: It was rumored that Bo was going to be out of town.

BODVARSSON: I leave at 4:30.

SAGÜÉS: Okay, very good. Although Bo Bodvarsson and Joe Farmer are not members of the roundtable discussion themselves, I think that it's very convenient that they're here in the audience, because periodically we may have to refer to some of their work.

And I'd like to start the discussion on a somewhat free format for right now. But I think that it would be very desirable to start with a discussion of the many comments the panel members would like to make on the models that we saw today that were presented by Bo Bodvarsson and Joe Farmer.

So what I would like to do at this moment is to open the panel for discussion for whoever would like to start making any comments.

EISENBERG: Could I ask a clarification? Are you asking about the models or about how well the models are good examples of validation exercises?

SAGÜÉS: I think that I wouldn't make any limitations at this moment. Just go ahead.

EISENBERG: I could make some comments about how well they might fit in with a validation approach. I guess I was a little disappointed in some of the examples. Bo Bodvarsson seemed to indicate that if--and I think Konikow should relate to this--if a calibrated model matches the data, that it's a demonstration--that seems to show that it's a proper calibration. It doesn't necessarily demonstrate validation, and yet it seemed to be portrayed as a validation exercise.

About Farmer, the Farmer examples, they show that the short-term measurement rates were confirmed, but it doesn't really respond to what may be the key question, which is can you extrapolate these data in these models over long times.

So I think in a sense, the questions that might be key are not answered. Can these models be extrapolated to long times and large distances, and how do we know? And is there assurance that alternative models with different implications for performance are not compatible with the data? What seems to have been shown is that the models that were proposed are compatible with the data. And what evidence is there that different processes don't arise over these long times and space scales?

And, finally, with the increased reliance on the waste package in EBS, have the models that support those components, has the support for those models been increased proportionately?

SAGÜÉS: Those issues apply equally to both models. By the way, more housekeeping, when any of the panel members speak, please say your last name first for those who keep records.

Do we have any comments on these statements on the part of members of the panel?

ORESKES: Oreskes, consultant. Yeah, I'd like to follow up and agree with that statement, and particularly with respect to the issue of the predictive accuracy of the calibrated model.

It seems to me that there's a conceptual confusion that takes place here, which is that it's a conflation of predictive accuracy with conceptual accuracy. It's extremely possible for a model to have a high degree of predictive accuracy, especially a calibrated model that's being used, as the cases we saw today were, over, as you point out, a specific time frame and a specific scale, specific geographic or temporal scale.

The fact that the calibrated model accurately predicts processes on that scale and time frame is no guarantee that it tells you that you have the accurate conceptual model.

Now, I don't mean to say that there's a simple answer to this question, because i don't think there is. I think it's an extremely difficult problem, and I'm not purporting to have an answer to it right now, but I think that this issue really has to be addressed, and I think there's a way in which when we call these things validation exercises, it seems to imply that the underlying process model, the underlying assumptions about what the processes are are valid, and I think that that implication, it seems to me, should raise concerns for us.

TSANG: Chin-Fu Tsang. I think there's definitely a difference between calibrated models and PA models. In calibrated models, you are looking at particular field experiments.

Now, the field experiment has a limited time frame, and you also have some features that you do not need at the PA model. For instance, when you do a pressure test, you have a high pressure gradient. For a PA model, you probably don't need such high pressure gradient near the well bore, and you say you have very important, in fact, near the injection point, in the PA model, you don't have to worry about that. That's one thing.

The second thing with calibration models is that if you calibrate, you can use a not so accurate model and hide a lot of things in the parameter value, which is fine for little short-term extrapolations. You're going to reproduce the next set of field experiments, that's fine. But you don't want to extrapolate to 10,000 years, 100,000 years, to a slightly different site with slightly different properties. You really have to be careful.

So I think that is a step to go from a calibrated model to the PA model. And one should handle that appropriately. They're not the same thing necessarily.

RUNNELLS: Runnells. I would just comment that Bo Bodvarsson was particularly careful I think to specify that his model as presented was for a particular site, a particular set of rocks, if you like, and a particular, I won't say time frame, but I think it was implied a time frame. There was no hint there that this was a generalization. So I think the fact that you can hide some of these unknowns, not hide, incorporate some of these unknowns into the parameters is somewhat acceptable when you specify, as he did, the model for this particular site, this particular time.

TSANG: I think the PA model is appropriate to hide some things, but you just have to be careful what to do when you're having such long-term predictions.

ANDREWS: This is Andrews. I think the issue has been raised about, but let's talk about the UZ flow, about predictive accuracy for the intended use of that particular model. The intended use, one intended use anyway, there's several others, is the average and spacial distribution of flux at repository horizon, of course something that's not directly observable. It's only inferable from some tests and from the model itself.

And I think what Bo showed first through a series of calibrations, and then through some, call them whatever you want to, confidence building, is that within a factor of two to five, perhaps a factor of ten, he could reasonably predict, and I'll use the word predict, the current present day percolation flux at the repository horizon. Coming at it from a lot of different angles, from temperatures, from chlorides, from strontium, from Chlorine-36, et cetera.

No one asked Bo to make that is the number 3.1 or 3.2. We asked is it between 3 and 10, or 30 and 100. That's the present day.

Now, it's also going to be used as a projection into the future, which requires some other forcing functions, in particular, climate change and the uncertainty in future states of climate, and future changes in infiltration that result from those future changes of climate. But as a starting point, if I just look at that one particular aspect of it, I would say that it has a very reasonable predictive accuracy for that particular aspect of the model.

APPLEGATE: Following up on that--Applegate, AGI--following up on that, I'm trying to think of it from a sort of policy maker's perspective, and again I'm hung up like a couple of the others are on this distinction between calibration and validation. It seems that at the heart of it, validation should be a reality check.

And the challenge here is that if you're viewing it as that, you're doing a reality check, and I guess the best way to put it is you're doing a reality check in Y2K, but the reality that you're actually trying to look at is Y12K.

And how do you get around that? How do you get around that problem, sort of getting beyond the calibration to the--in other words, the danger is that you're promising too much in terms of even describing it as validation in that context.

KONIKOW: Konikow. I'd like to say a few words. I don't have any particular criticisms or comments on the specific models that were used, but again, what I heard yesterday and particularly today was what I interpret as a lot of wordsmithing and spin doctoring related to the concept and terminology of model validation.

I was really kind of surprised and maybe even chagrined at how ingrained and pervasive within the small community related to high level repositories this concept and desire to validate models is. It's even on the cover sheets for reports that Dan sent me a couple days ago, even a check-off box for model validation. And this really amazes me.

It's something to check off. We've done it. And one of the dangers of course in doing this is that--well, there's several dangers. One is that you imply models can indeed be validated. Another is that you imply, and a lot of people take this implication that once the model has been validated, there's no need for further testing, because we have valid models.

If I look in this particular report that was sent to me, again I just keep seeing self-inconsistencies dealing with this whole concept of model validation. And, again, I'm not criticizing the model itself or what was done for model testing. But in the section on model validation, it says this model cannot be validated vigorously. Okay? And so every once in a while we see a hint that this really can't be done. And they say, however, it can be partly validated, whatever that means. And again, this gets into the whole concept of what it means and how different people interpret the terminology.

This morning, we heard basically it's a gray scale, that there's a continued gradation of degrees of validation because you define the term to mean confidence. I think the term validation and the concept of model validation to most people, to scientists and to the public, is a yes, no, statistics. You validated it or it's not valid.

If we look again on Figure 21 from this particular report, I found it interesting an illustration of the validation tests show four particular tests, and he describes the criteria, you know, expecting the validation to be successful if the data lie within the 95 per cent error calculated by the model. And then two of the four tests, the observations lie outside the 95 per cent confidence interval. And so the implication made in the report is not that this invalidates the model. The implication is that we've only partly validated it.

Well, I just--you know, I just don't buy that. It just seems--I don't understand why you're so hung up on using validation. I have my suspicions. But I think the whole concept of model validation as you're using it is invalid.

SAGÜÉS: Since this is a roundtable discussion, we'll for the time being, we'll limit the discussion to a roundtable. I guess Linda Lehman has something to say at this moment.

LEHMAN: Yes, Linda Lehman, Nevada. Lenny, I think a lot of this goes way back to the days of early NRC regulatory development when in Part 60, we were looking for some assurance that the models were at lease consistent and correct.

However, over time, and after being involved with the INTRAVAL process for six years, I've kind of come to the conclusion that I don't think it can be done. And some of the experience in INTRAVAL, for example with Yucca Mountain, we actually had a Yucca Mountain test case, and in that test case, most of the participants used one dimensional matrix flow model. I used a two dimensional fracture flow model, and our challenge was to predict saturations in a deep borehole based on some shallow borehole data.

Well, some of the models predicted part of the curve better than others, and for example, maybe mine predicted the upper part of the curve best, and the matrix flow ones predicted the lower part of the curve. Well, then the INTRAVAL went through this whole process to try to figure out which one was better, and they couldn't do it.

Yet while we could all do a reasonable job in matching the saturations, the velocities were really, really different. We would get velocities which ranged--or flux rates, I guess we were looking at, from .01 millimeters per year to 7 or 8 millimeters per year, and still match fairly well the saturations. So that led me to conclude that we have to look at more parameters when we are trying to, as I say, validate.

Now, what I've come up with is that we can't validate, but that we can build confidence, and the way to do it is somewhat different I think than the validation approach that was presented today, you know, confirming that the models are numerically correct, and assuring the data inputs are okay. I think it's something more basic than that, and it's something that Bo did in his models, basically used all the data sets that are available.

For example, I'm going to use the example of the saturated zone. I have developed a fracture flow model, whereas up until recently, everyone was working with basically matrix flow models. I was able to match temperature and pressure at the water table surface.

The Department of Energy has only tried to match potentiometric surface, and you can match that potentiometric surface in a whole lot of ways, but you can't match the potentiometric surface and the temperature profiles as many ways.

So, to me, the key word is lets constrain the results. We have solution; we need to constrain it. So let's go about constraining it in the best way that we can. And we have other data sets we can use. We have vertical head distributions which aren't being used. We have temperature and we have chemistry.

And I think as a first step in building confidence in the model, and true we can't extrapolate it, but at least if we could get some confidence that the underlying concepts are correct through matching these other data sets, then I think that goes a long way in assuring the public that we have something that we can go with.

PARIZEK: Parizek, Board. The unsaturated zone study is somewhat unique in terms of the effort that's gone into that. So of the data sets, what else could you have? I mean, here you had the perched water. You had various gas compositions. There was the age dates of the water, and so on. It's kind of unique to have that much to work with.

What was not mentioned is really like the vein development, cement materials in the mountain, which over the long geological periods of time, say, well how much water would have to go in there, some of the U. S. Geological Survey work that's saying over the years, you have to have this much mass of water to deposit those minerals.

So it's sort of like an analog for the models. You know, if the models are not way off because of the geological observations you make, you feel good. So I'd keep asking, well, where is the analog support? That gives you some other way of underpinning the concept. It's sort of like what Zel Peterman did at the Beatty meeting for your discussion. You had a suggestion of the pattern of flow, and the mass of geochemistry data, such as it exists, good or bad, supports it. It doesn't argue against it. So that's another line of evidence, and so on.

So we need to have for a complex system like this as many different observations as you could make from the different disciplines that help support and help build confidence in the conceptual model that you've got. That's probably as good as you're going to be able to do.

And then that brings up the audit or the post-audit things, Lenny, which you could probably comment on as to how good are we on audits. But that's really observations you make after you make a prediction, after you do some engineering decisions, to see if it's performing like you've predicted.

And maybe the best chance for Yucca Mountain is to begin putting wastes underground with the idea you're going to be making observations while you do that to see if everything is working, and you don't close the door, and the longer the door stays open, the more chance we have to get those observations, which is not really--it can be misunderstood. The public might say that's because you guys really don't know anything about the mountain, or you don't ever intend to take the waste out of the mountain. We don't trust you.

Where on the other hand, we say no, we want to ventilate it, we want to keep it cool, leave it there, but if you find out there's something wrong with it based on the actual observation of how this thing is performing, you have to trust us to do something about in a reasonable time period rather than slamming the door two days later and say we can't touch it ever again.

So this idea of a post-decision audit is sort of like that, and for Yucca Mountain for 10,000 years, what kind of audits could we conduct, you know, is always the concern the public would have. But maybe some comments on audits and how good they are or how bad they are, just from a physical flow or chemical transport models would give us a sense of where you're coming from.

KONIKOW: Konikow. I've conducted a number of post-audits, and what these are basically is looking at the true predictive accuracy of deterministic groundwater models of various types. And what I mean by true predictive accuracy is that we've gone in years after the predictions were made to see what the outcome is, and I've published a number of papers on this, and in general, for models that were very well calibrated for periods ranging from ten years to forty years, making predictions of several to ten or twenty years into the future now that the deterministic models have been around for a number of years, we go back in and see how good the accuracy was.

And in general, the predictive accuracy was pretty poor, not very good. It was variable and there were a number of reasons. Some of the reasons were, and I think a lot of the reasons have transfer value to the Yucca Mountain situation, some of the reasons were that the predictions of future stresses were not very accurate. Some of the problems were that single predictions were made rather than evaluating a range of uncertainty in the input. And that's a mistake that we tend not to make any more.

So in a sense, the prediction that was made really should have had confidence bounds around it and it didn't. And so one of the interesting things, we'd go back and see what those error bands would look like, and see if the predictive outcome really fell within that or not. But just looking at the actual prediction and comparing it to the observed, there are very significant errors. And so at least in some of the cases, I would predict it would have been outside the confidence intervals.

Other reasons were that there were conceptual errors in the model, and of course other reasons were there were errors in the parameters, in the estimates of parameters, that on a short-term prediction and during the calibration, did not show up, or the match was not sensitive for the calibration period or the history match, or as was mentioned, compensating errors were built into the parameters. That doesn't show up until you make a longer term prediction and see what's going on.

Another possibility, and I think this was true in some cases, that the conceptual model was weak, and it may have been okay for the history matching phase, but then when you got into prediction under either a different set of stresses or a longer time period, that conceptual model just was no longer applicable.

In some cases, it was as simple as using a two dimensional model when they should have been using a three dimensional model. So the record really isn't that good, and this is for periods of, you know, predictions on the orders of years to maybe decades, and we're talking about 10,000 years, and this raises concerns. And, again, it gets to, you know, when you say the model is validated, what does that imply in terms of long-term predictive accuracy. Because even in the performance assessment framework, in this probabilistic framework, you're still using these underlying deterministic models to make the predictions.

SAGÜÉS: Very good. Applegate, and Tsang.

TSANG: Tsang. I think a lot of the issues that has been mentioned have been considered in the nuclear waste community in the process of worrying about validation.

One very good example which I very much recommended is the SKI '94 Report that's published by SKI in 1997. It is the SKI's performance assessment exercise in which they look very carefully at all the FEPs, features, events and processes, and get the experts to have an elicitation of the events, and what they call process importance impact diagram.

I have two viewgraphs. Should I show that to you to show the results? And it has a very good discussion of uncertainties and errors and relationships, so I think that is a report everyone should read.

This is one example in which they look at the conceptual models of different fracture rocks. So the three groups at varied--different conceptual models. And then they try to get the results and errors involved. And this is a picture I think that's quite interesting. Taking a model like Lenny was saying, all the predictions must have an uncertainty range, and I think that's a very important quality.

Think of prediction as--you have evaluate how much confidence you have. This uncertainty range is different from how confident you are of the results.

When you have a big uncertainty range, you have a high confidence it's within the flow, porosity, within zero and--well, it should be between zero and--much improvement in your range. Again, you have confidence. So I think the range, the uncertainty range and confidence are two different objects.

Here, they use three different models, which are completely different, discrete fractures, stochastic continuum, and simple models. And the range of errors is quite different, and so they look at the whole thing to do this kind of performance assessment.

So I think we're addressing some of your concerns. And, of course, the question of--is also important.

APPLEGATE: I'm very glad this issue of post-audits and monitoring has come up, because they seem absolutely critical to the notion of validation.

But they also point out what I think is the single difference between, and this has been talked about a bit over the last two days, between the license application, the LA, and the actual decision by the President about site suitability. And essentially, the difference being that the LA is a regulatory decision and we've got to recognize that the other, the SR, I guess, is a political decision.

And whereas, I think the monitoring has to be absolutely a fundamental part of a license application and should be recognized as part of validation, it's of virtually no use in terms of the political decision.

And the only thing I'm going to try to equate this in with the, since we've been using airplane analogies here, from a political standpoint, assuming that we've decided the SR would be deciding that we're going to get on this airplane, the notion that monitoring was of any value from a political standpoint would be that there were indeed parachutes on this plane. However, the situation being that nobody has ever used them and nobody has any confidence that they really would work, and that the politicians certainly would feel that once you put something in the ground, it's not coming back out, and that's been universal in these types of situations.

EISENBERG: Eisenberg from NRC. I'd like to respond to Konikow. I want to make sure we don't get all wrapped up in a semantic argument. From the negativist point of view of scientific theory, validation is not possible. All scientific knowledge is tentative, subject to the next experiment, which could overthrow all the principles that everybody has agreed to up until that point.

However, from the positivist point of view, confidence in the models is raised by a variety of testing activities, some of which have been discussed today. We have to remember I think that the purpose of this whole program is not to make progress in science. We may have to do so in order to get where we need to go, but the purpose of the program is to make an important national decision. And from that point of view, it's appropriate to use these positivist techniques, these confidence building activities, and the fact that this community has chosen to sometimes call them validation activities I think is not such a bad thing.

I should mention that number one in this White Paper on model validation produced jointly by NRC and SKI, we do say that the terms confidence building and validation are used interchangeably. I'm sure that's not acceptable in some circles, but they are--I think what is intended is confidence building in a strict semantic sense.

And also, the scientific community, I was at a meeting of the GEOTRAP study in June, and one of the conclusions is is that the whole international community concerns with waste management has come to the realization that perhaps confidence building is a more appropriate term and is a more appropriate goal for these programs.

NEWMAN: Can I say a word about that in the context of the waste package? I think it was decided a number of years ago in several countries, and I'm not sure if the U. S. really comes into this category or not, but that you never had any chance of validating a model that was associated with the initiation of extremely rare corrosion events, such as pits. I use the word rare simply in a geometrical sense. That is there are ten to the nine axioms on every square meter and any one of them initiates a pit each year. So that's one in every ten to the 27 axioms per second initiates a corrosion event.

And I think those of us who thought about that really don't have any desire to get involved in validating models like that, although we recognize that if you want to answer questions like how many holes is it going to be in the container after 1,000 years, you might have to get into that.

But since you've made this decision to use this very expensive material, that means you have the opportunity to have another much simpler kind of validation, which is simply to show that even if corrosion--even if you force the corrosion to start, it will in fact stop. And that's a much easier kind of--or what I call an arrest criterion is a much easier kind of approach from the point of view of prediction and can be validated much more easily, because it essentially converts what is a classically stochastic kind of problem, that of localized corrosion, into a deterministic one. Namely, if you're lucky, you'll show that under all the conditions that are relevant to your repository, even if you force the corrosion to start by temporarily increasing the temperature or the chloride or something, when you bring the conditions back to the real conditions, it will stop.

I think that's the only--just speaking from the waste package corrosion, that's actually the only kind of model that you have any chance of validating, is an arrest model. Now, you might be unlucky. You might find that under some of the conditions that you've got, if you do that, the crevices will carry on corroding under a condition that you can imagine existing in the repository. Then you have to go back to an initiation type philosophy. And good luck.

ORESKES: I wanted to make a point about the issue of the scientific knowledge and validation in a sort of larger scheme of things.

It seems to me that what we're involved in here is quite different actually from what goes on in science generally, or what has historically gone on in science, which is that we're trying to make a decision here by a certain date, and it's extremely admirable in the history of science for scientists to have a date that they have to solve a problem by. And so there's a kind of anomaly about this that I think we shouldn't gloss over, and it's not to say that

that's a bad thing. I mean, it may be perfectly legitimate from a social and political point of view to say we have a problem and we want to do the best we can with the available knowledge.

But that's really different than a situation in which over the course of time, a scientific community comes to a consensus about an intellectual question, and I think it's really different in a way that I think it's important for this Board to, I hope, to think about. I hope that you'll think about it. Which is that it seems to me that one of the things that we know almost certainly in this sea of uncertainty about nuclear waste is that there will be significant changes in scientific knowledge and technical capacity in the course of the next 10,000 years. I think that's, as a historian, one of the few things that I would feel safe about predicting about the future.

I mean, if it passes any kind of guide at all, we can expect even 100 years from now, much less a thousand or 10,000, we will hopefully know so much more about so many of these questions. So that's where I'm an optimist about scientific knowledge. And I think that the really--one of the really important things about that insight is that we have the capacity to make future modifications and adjustments through monitoring, and to make improvements as we learn more about this problem in the future.

What worries me about the language of validation or even confidence is that to me it doesn't seem to invite a kind of deep appreciation of the fact that this possibility for improvement could take place in the future. And I'm not talking so much about among scientists, because I think among the scientific community, we all do science or we're involved in science because we have the hope of improved knowledge in the future. But I'm thinking more about when this gets transmitted into a political arena.

It seems to me very important for the Department of Energy and for this Board to, when the site recommendation goes forward, to do it in such a way that reminds the political community that there is a future task ahead that involves learning, monitoring and modification, and that that future task of monitoring and modification is every bit as important, if not more important, than the work that we've done to date.

And I know that this is something that people in this room know, and I don't mean to imply by any stretch of the imagination that people here don't know this, but when people talk about validation and when they talk about valid models, I think to most people outside of this room, as many have said, I think most people think that means that we know what's going on. And so I would just really like to strongly say that I think the language that we use is terribly important in terms of the message that we convey about what happens, not just in OO but in 50 and 100 and 200 and 500, and that that's part of what I think the issue is that we're facing here now.

LEHMAN: Linda Lehman, Nevada. I think a lot of the problem has to do with expectations. I think there are a lot of differing expectations on the word validation or confidence building. For example, I think the public when they want to see the results of a performance assessment, yields a dose, they want to be sure that that dose is lower than some standard.

I think some of us modelers have done a lot of modelling. Our expectation is, well, I don't have a lot of confidence in this result, but if I've done a lot of testing and a lot of comparisons, a lot of calibrations like Bo has, well, then I have a little more confidence that maybe my model is better. But I wouldn't be willing to stake my life on it.

Maybe some other program participants have a higher expectation of what they're going to get out of it. I think basically what the program is using it for is a decision document or a number to make some decision on. And I think these differing expectations, especially like you say, the reaction to the word valid means that it's real and it is very real to the members of the public, but maybe to Norm or

Tim McCarten, it's not a real number, but it's a realization.

So I think that needs to be conveyed.

KONIKOW: Konikow, USGS. I'd like to agree with Linda and with Naomi, and I think, contrary to what Eisenberg said, I would argue that it is more than a semantic issue, that there are some real substantive issues here, scientific and otherwise.

I'd like to reiterate what Naomi said, is that the term valid has a certain meaning to most of the public, and it carries with it an aura of correctness that I think most modelers would agree is not really there. And I think one of the ways, one way to look at this in terms of what's the implications, why is this a problem, straying a little bit from science, I would recognize or just, you know, state that, maybe you're not aware of it, but DOE does have a little bit of an image problem. In all circles, DOE does not have the greatest reputation for being straightforward and honest and reliable. And, I mean, I trust you, but not everybody does.

So the problem with this focus, and really today harping on model validation, what concerns me is that you're not using the same definition that everyone else is. And, you know, if I think back to reading Alice in Wonderland, you know, the Red Queen, I believe it was, decided that terms would mean whatever she meant it to mean, whenever she used them, and it wasn't necessary and she could change the meaning at will. Well, you know, she came off as being silly, and as being nonsense.

Very recently, there's a widely publicized case in which a famous world leader made some statements about his personal life based on a definition of a term that was very different from what the public took as the meaning for that term. And the consequence of that is that he came off being perceived as dishonest.

And what I see here in DOE, with a high level rad waste community, continuing to harp on model validation is that you're going to come off as being either silly or just dishonest by implying an aura of correctness to the models and reliability to the models that is just not there.

One of the real dangers of that, when these things go to court, which is a distinct possibility, you are opening yourself up to attack on the issue of validation. You are opening yourself up to attack on is this model really valid? You said it was valid. Is it really valid? And you're going to get mired down in all kinds of critiques on how valid that model is and whether or not it's really validated, what it means, and you're going to say, well, we didn't mean that as a valid model. We meant there was confidence. We have confidence in the model.

Well, if you have confidence in the model and that's what you mean, why don't you say that? If you mean the model has been well calibrated, don't say it's been validated. Say it's been well calibrated.

What are you trying to gain or who are you trying to impress or what are you trying to prove by saying it's validated when you've defined this to mean something different than what everyone else seems to think that this term means. I'm not sure what your goal is in continuing to use this term validation that means different things. And when you get to the political decisions and you explain to the politicians that our analyses are based on valid models, are you going to clearly tell them what you mean by valid, or are you just going to say these models have all been validated? Are they going to know what you mean when you say that it's all based on valid models?

When you're going to get challenged in court on these things, what it's going to do, among other things, is divert attention away from the true substantive issues and how good the models are and how good the predictions are, and you're going to get mired down in nonsense. But it's going to make you look bad.

SAGÜÉS: I made a note here to maybe ask Dr. Andrews in a minute, since he did present a couple of definitions of validation on the transparencies, and it looks to me like we are discussion quite a bit the meaning of a word, and maybe we're wanting--many of the items that you mentioned presumably would be solved with an adequate definition.

KONIKOW: Not if that definition is different from how people perceive it.

SAGÜÉS: Or maybe a different definition. But perhaps what I'm going to do is I would like to invite Dr. Andrews to perhaps address some of those issues, and then anyone else if you have some comments.

ANDREWS: Okay, thank you. I think we have to be careful. That word probably means different things to different people. I bet everybody in this room would come up with a different definition of the word validity. If one said it was a reasonable representation because it is a model that we're talking about, it's not a reality per se, we will never test every square centimeter of the rock, or every square millimeter of every package that may be made, so you have to have an approximation, i.e. a model that represents as close as you can to "reality."

As Lenny pointed out, there's a number in historically models based on limited information that perhaps when actually stressed, didn't explain exactly, however you want to define exactly, the assessment of contaminant migration, or whatever aspects he was looking at. I mean, it was water, not contaminants. It was oil, not water or contaminants. A lot of assessments, a lot of models of all of those processes are created.

So I think if we say it's the reasonableness and the reasonableness is, I think Linda had a very good observation of the more independent lines of evidence that one can bring to bear on that particular process as it is implemented for the intended purpose of making an assessment, a prediction, if you will, of future behavior, the more independent lines of evidence that can be brought to bear so it's not just potential measurements, it's temperatures and chemistries, et cetera, the closer, the better chance you have of it being a reasonable representation.

Is it unique? Probably not. And the non-uniqueness of those models are addressed. They have to be addressed to evaluate these key decisions. And I would argue that in science and engineering, those key decisions happen all the time, and in lots of cases, they are driven by a schedule. Building a dam or putting up a power plant or putting up a bridge across a road, they're driving by in some cases schedules, and they are based on scientific observations and models in many cases.

So can they be improved? Yes. Will they be improved? Assuming the project goes forward, yes. I mean, the improvements in each of these aspects of science are to be expected. There's plans in place for those. Are they valid in the traditional sense of the word? Probably not.

But are they adequate for the intended purposes? Probably so, with the uncertainty hopefully captured in a reasonable fashion.

So the decision makers who have to make decisions know what the uncertainty in certain of these aspects are.

ORESKES: Can I asked a question, though? Then why don't you just say that the model has been tested and found to be adequate for the available purpose? I mean--

ANDREWS: We probably will.

ORESKES: Well, no, but I was listening today and I was asking myself the question when people use the word validated, could you substitute the word tested? Could you say--I mean, in every single case, it seemed to me that you could, and then that raised to me the question of why you didn't say that. Because it seems to me that using the word tested would have a much more transparent meaning to most people in the scientific community and in the general public.

ANDREWS: The TRB wanted this discussion of validation.

EISENBERG: Can I just jump in for just a second? Most of the models will not be tested in a direct fashion over the time periods and spacial scales of interest.

ORESKES: But they're not being validated over the time scales and spacial scales either.

EISENBERG: Absolutely not.

ORESKES: I mean, all tests are partial tests; right?

We always test pieces of things. We can never test the whole thing. But it seems to me that what you're doing are tests, and I think that--I don't think there's anyone in this room who would imply that the tests that have been done aren't good tests, or there hasn't been a lot of good work done to support these models. I think it's very clear from the presentations there's been a tremendous amount of really good work. But the question is what you take away from that work and how you present it, and I think those are the issues that people outside DOE are concerned about.

SAGÜÉS: Debra has some questions, and then I would like to steer the conversation after your comments into something perhaps a little more concrete.

RUNNELLS: Something Naomi said triggered this, and that is the schedule driven science. In my academic life in 30 years or so, the schedule is not nearly as important as it is now, when we have scheduled deadlines we have to meet. We think we do pretty good science and engineering. We still have to meet those deadlines.

Now, the work--when I say we, the work that I do that we--that my group does is similar in some ways to Yucca Mountain. We deal mainly with mines, and mainly with mines in Nevada. Those mines have the potential to do a couple of things. One is to seriously alter the hydrologic regime. These are large open pit mines. And they have a very great potential to contaminate groundwater with metals primarily.

We use the same models, the same sorts of models we've heard described here today for hydrology and for geochemistry. But there's a profound difference, and sitting here, I finally identified the difference between what we're talking about with nuclear waste and what I do every day with other contaminants in a similar environment, and that difference is that we recognize the impossibility of predicting some of these things. We and the regulators with whom we deal, the Bureau of Land Management, the Forest Service, the state regulators, recognize that we cannot predict and we all admit it, we cannot predict the chemistry of a pit lake in an abandoned mine 2000 or 3000 or 4000 years from now.

We cannot predict adequately the impact on the groundwater regime of an open pit mine a mile long with all of the complications that go into that fault, even so on and so forth, with the recharge of water. As a result, we have a contingency plan. We will predict as best we can what will happen on a short time scale. For that, I mean less than 100 years, and more generally, ten years. And what if we're wrong? Everybody has to understand that we may be wrong, even on a time scale of ten years.

I won't call it an agreement, but the understanding that has developed is that we will cover that with intensive monitoring, exactly what you said, Naomi, also about the monitoring. Having recognized the impossibility of predicting 5000 years into the future the chemistry of a lake, we will monitor the chemistry of that lake, and if we see it deviating from our predictions, and this is I think also different than Yucca Mountain, we have a contingency plan.

What if it deviates, what if something goes wrong? What if instead of the water being good quality and supporting wild life, suppose it's loaded with arsenic, then what will be do? And the regulatory agencies with whom we work require two things. They require the monitoring plan, and they require the contingency plan, so that if something goes wrong, we have some backup plan.

Now, sitting and listening now for a year or so to discussions of Yucca Mountain, I'm not sure that we have a backup plan. I'm not sure we have the second half of the activity of the agreement, or the understanding that allows a very difficult scientific problem to be accepted by regulators, the scientific problem being contamination of groundwater in a water poor state, Nevada, and hydrologic modelling that's difficult to do.

So I would--I don't have an answer. I'm not even sure I have a question, other than isn't there some contingency plan that could be discussed, outlined such that the public and the regulators have some level of comfort that if the predictions are wrong, that positive action can be taken.

The retrievability, I've heard that mentioned occasionally, retrievability is a sort of contingency plan. But I don't often hear that, if ever, discussed in our discussions recently about Yucca Mountain. But in this other world, that contingency plan is absolutely required, because we recognize the weakness of the predictive modelling period.

SAGÜÉS: Okay, a very important observation. Now, if we could continue in this vein, especially with this new area you just mentioned, Don, but I would like to at least for a little bit to go to perhaps more specific issues.

I think that this may be a good time, and some of you may have quite a bit to say. Today we heard an example of a model prediction that may have a great impact on what may be expected to happen in the mountain. We heard that a 1000 millimeter per year percolation flux threshold for seeping. Now, granted, that that was presented as a preliminary type of observation, but certainly the kind of things that models, if validated, would change very much the way in which we would look at the mountain.

Do we have here within the panel any specific comments about that kind of number? Maybe some members of the panel may have something more to say.

PARIZEK: Parizek, Board. When there was a comment earlier in the afternoon, there was a question that didn't get asked, and it really could have been directed toward Bo, and I think he's since left, but--

SAGÜÉS: He's right there.

PARIZEK: Good. Earlier, in fact, we asked earlier about the shape of the tunnel, and the idea, as an example, if it's a perfectly round little tunnel, maybe the water will weep down the sides and there will never been drips, even though water enters the tunnel.

On the other hand, if you have an irregular tunnel, because its roof collapsed, and so on, then maybe water has a tendency to want to hang up in the irregularities in the roof, and it will drip.

So here's a case where no matter how good the models were, unless you know whether it will drip or not, and what conditions may give rise to drips, maybe that 1000 millimeter number has some limits to it, because of the special condition of the shape of the tunnel, because it's dynamically changing in time.

So, Bo, do we have anything specific about tunnel shape and stability? And if you start rattling the roof down and you have, you know, ragged roofs, will water hang up and want to come in on your head, versus a round tunnel?

BODVARSSON: Bo Bodvarsson, M&O. Your question is a very good one. We started seepage testing two years ago, so it's a very young program and a very important program. As a part of that, we identified several things that need to be looked at. One is certainly the approximation of a continuum model for a discrete fractured site, and that's one thing we want to do, is to evaluate the results from a discrete fracture model.

The other thing is the size of the opening, and the changes in the size and shape of the opening. The size and shape of the opening, Chin-Fu Tsang, which is right there, is doing the PA seepage model for Bob Andrews, and as a part of that work scope, is to change the shape of the tunnel based on an AMR that comes from the EPS that tells us how they think the shape is going to change.

In addition to that, we want to do laboratory studies where we can actually control the shape of the opening, which is much easier to do than to drill a square niche, which is not easy to do. So we are addressing that issue.

Preliminary results that Chin-Fu and his co-workers have gotten, and they can explain it later in detail, based on what they've gotten so far, we don't see a lot of difference between those examples and the regular smooth niche. But that's subject to verification.

Finally, since I have to go, I want to make--can I make a couple of comments?

I really agree with all of what has been said in terms of the validation should not be used for our models. And I couldn't agree more with that because I think it's always going to get us in trouble, and we don't need to use it, unless NRC tells us we have to use it, and then I'm going to back off. But if we have a choice and we can say confidence building in the model, and we can do the same thing with it this afternoon, show the public all these different data sets independently, I think we'll give them a warm and fuzzy feeling. So perhaps we don't need to use that word.

And I think the main argument has been over that word rather than the approaches, and you correct me if I'm wrong. So that's all I wanted to say. Thanks.

SAGÜÉS: Very good.

TSANG: Maybe let me add a few more words about seepage modelling. We look at a calibration model, we look at the parameters very carefully, because the field experiment, you have a lot of trenching effect, which is probably not needed in the PA model, and also it has a point source. And so we take those into account.

We look at the shape dependence quite carefully, especially the mechanidate plat, and I review over the calculations for the mechanical degradation, changing permeability and rock fault, some of the work done by the disturbed zone group. It's quite interesting. The keep lock theory was used to make the calculation on the one hand, which showed the rock fault occurs something like once every hundred meters, of that order.

In that case, you only need to worry about one rock fault at the same time, and the cavity, a hole there does not create extra accumulation of moisture. So it does not affect the results very much.

Then the other way is to do a redax calculation where the fracture opens. So we're looking at that very carefully. It turns out that in many cases the vertical fractures get closed, and the tangential fracture opens more in many cases, in which case actually it's better for seepage. That means that there's a better chance for it to go around the drift. So all these are being evaluated and we try to look at the uncertainty range, and that kind of thing.

Now, just for the--many were asking what model has been invalidated earlier. I was just thinking in terms of seepage model, I can say we have invalidated John Phillips model, we have invalidated Calvin's relationship, and we've probably invalidated hydrology. Let me explain.

Number one, John Phillips model, as you know, he published a lot of papers on underground cavity seeping into it, and he mainly--we show that using his model, the estimate for seepage is two orders of magnitude. The reason is quite simple, because he used homogeneous flow, and whereas if you look at the heterogeneous flow, there is a channelling effect that what is more likely to accumulate, and the result is two orders of magnitude difference, which if you look at niche test, certain--does not work.

The second one, Calvin's relationship mainly says that you have a ventilated drift. The ventilation causes a big suction from the rock, and this suction is huge, capillary suction because of ventilation. And the niche test says no, it is a capillary barrier with suction, probably because of low--effect. So we have to use a capillary barrier concept.

And then why does the hydrology doesn't work is because you have to worry about hydromechanical effect. Once you have exurbation, the Joe Lenz measurements show that the permeability increases by two orders of magnitude on the average, and that turns out we have to take that into account, and that also is the reason the alpha value, the van Genuchten alpha value is different by a factor of 100, two orders of magnitude.

So there is a difference between regional alpha and the niche scale alpha, but the niche scale alpha is what is controlling seepage. So we did try to invalidate something like this.

SAGÜÉS: Let me make a comment. Again, the validity of this kind of model, since we are not taking into consideration the fact that that rock is going to be heated to a fairly high temperature for hundreds if not thousands of years, wouldn't that throw just about any modelling effort just out the window?

TSANG: We did look into the thermal problem, and I'm interested in coupled thermal hydromechanical. It turns out the thermal problem at the current plan, you will dry up the near rock, the near field within, say, half a meter, it would dry up.

In that case, as far as water flow goes, is that should get better, because the--goes down, the fracture permeability goes down, and the water is harder to flow into the rock. It tends to go around. And then if you look at thermohydrological a bit more, away from the niche, about five meters away, there is what's called reflux zone, boiling and condensation and evaporation. There, that could be the silica deposit deposition and the permeability would go down. And that is like a shield.

But this is just rough discussion right now. We are looking at the THC calculation, thermohydrochemical calculation, looking at the impact. So we are looking at the problem and hopefully we'll have some results this time next year.

LEHMAN: Yes. Chin-Fu, I don't know if you saw this presentation, but Dr. Parizek and I were at an NRC technical exchange a few months back in San Antonio, and there was a woman, I believe her name was Deborah Houston, who looked at the shape of the tunnel and what she did instead of using a smooth tunnel surface, she actually used a sine function across the top. And so by varying the sine function, she felt that she was getting three orders of magnitude more infiltration with that type of shape, which she thought could be expected over time, than with the smooth wall.

So I don't know if you're aware of that work or if it's a disconnect, but it would be interesting to resolve.

TSANG: I'd be interested to look at that and resolve that.

SAGÜÉS: Taking advantage of this. I would like to take the conversation over a little bit to materials performance issues, and I wanted to express something that I have mentioned before, one of my main concerns, but it has to do a little bit with what Dave indicated earlier. And that is the fact that we are not only having to deal with a model that may or may not be appropriate, to use a different word this time, but rather, it's that that model has to be appropriate over an extremely small time frame.

In the case of materials performance, we have--or specifically corrosion--we have two issues. One could divide the program into two issues. Issue Number One is is there any viciously fast mode of corrosion that will create a problem in a very short time?

For example, pitting corrosion, crevice corrosion, stress corrosion cracking, and the light. And much of the effort until now has been devoted to determining how likely those fast modes of deterioration will be. And, indeed, Dr. Newman just suggested one approach that is somewhat different from what has been used most of the time in the project.

However, even after you solve that problem, now you have the question as to whether there's lower forms of corrosion, specifically, for example, passive dissolution of the metal, are going to be the kind of things that one can rely upon for extremely long-term durability. That means that the system as we were discussing earlier today has to survive at the rate of corrosion that is going to be on the order of, say, one-tenth of a micrometer per year for periods of time that will be at least 10,000 years, but one would be more comfortable with perhaps 100,000 years, because one wants to have the medium of the distribution of damage safely away from the goal that one wants to achieve.

Now, we're going to be relying in this particular repository on one concept, and that is the concept of metal passivity to provide the material durability. We're not relying on, for example, very slow active dissolution of the metal, as what would be happening if we have, say, just plain steel environment.

Here, we are dealing on the formation of a very thin layer that barring these very fast modes of deterioration, is going to have to stay put, and chewing through the metal very, very, very slowing over a 10 millennium, if not 100 millennium at least.

Now, there is one problem, and that is that this passivity trick that we'll use enough for a whole bunch of high performance alloys, this has really been in use for the protection of engineering materials for about 100 years. I would say the Twentieth Century in real application. The phenomenon was known some time early in the Nineteenth Century. But nevertheless, we have here basically 100 years of known performance, but we have 100 times 100 years of performance, but we really want perhaps a 100,000 on the average, as I said before, so in here with an extrapolation gap, if you will, there's going to be an extrapolation gap of about three orders of magnitude of known performance.

And the question I would like to bring up right now is in how many instances do we have in the history of science, the history of engineering, situations in which we have had to extrapolate so far in advance beyond proven engineering, a ground tooth performance. How about Newton's Apple and rockets to Mars?

SAGÜÉS: Okay, explain that a little bit more. Okay, what has extrapolation got from Newton's Apply to rockets, interplanetary travel? It's a distance extrapolation.

TSANG: Well, it's really not my field. But let me try to say something. It is of course terribly impressive to me, Newton had the apple, found the gravity, and the rocket theory reaction, and then you can send the rocket to the moon, to Mars with terrible accuracy. I mean, that's just totally amazing. And this means that you really have to get the basic physics and chemistry right.

And so that's the reason I'm very hesitant about using calibrated models blindly. You need model calibration, no question about that. And you need model testing. But you need to understand the basic physics and chemistry processes and get the most up to date signs from the scientific community. Then you can do the best job you could about that. There's no other choice.

So then--and you cannot do better than that on principle. So the question then is that so I define validation more than just testing. Validation, you could do testing, plus understanding the processes, plus confidence building. So one can use those words.

But anyway, so I think the trick to the whole thing is, in my view, is how do you bring a maximum state of knowledge into this game. That is not so easy when you consider it. But anyway, that is all I can say.

SAGÜÉS: I guess the question is we'll do the best we can. Of course the question is is the best good enough.

ORESKES: If I could just follow up? I think the extrapolation gap is enormous, and I don't think there are any examples in the history of science or engineering that are comparable, and if anybody knows of any, then I'd love to hear them. And I think that's one of the challenges that we're facing here.

I think what we're trying to do here is unprecedented, and that's one of the reasons why I think it's terribly important for us to think about how we incorporate mechanisms to bring the latest state of the art scientific knowledge into the process, not just right at this moment, although it's obviously really important right now, but also continuing into the future. And I think it does require some new strategies.

NEWMAN: With regard to the particular thing that you mentioned, once again, I think the way to look at it is to try to speed it up in the beginning, and to try to create whatever the unusual surface conditions are that you might be able to anticipate in an accelerated manner, and then relax the system back to the real surface conditions and see if you've changed the way that it behaves in any way.

For example, some of these corrosion product layers that you mentioned may be ion selective. They may have a membrane property. So they might let the chloride ions in, but not be very good at letting the metal ions out.

One can create such a layer in an accelerated manner, and then examine its effects on the process. That's indirect. I'd have to explain in court how I could extrapolate from that observation to a guaranteed immunity of a nuclear waste canister. But that's part of the process I think of understanding, is that you have to have imagination and you have to be able to imagine all the things that could go wrong, and if you're not clever enough, you might miss one. But if you can think of all the scenarios in which this corrosion rate could gradually speed up with time or could become unacceptable, I think it's normally, at least for these cases, possible to simulate that in a short period of time, and then examine what happens.

I just wanted to point out one thing, since I'm only here for one day, and that's all passive films on chromium containing alloys are the same. You shouldn't come away with the idea that the passive film on Alloy-22 is different or better than the passive film on 304 or 316 stainless steel. It isn't. It's the metal that's different.

FARMER: I want to take exception to that. We've done x-ray photomicron spectroscopy and depth, and the film actually is different on Alloy 22, depending upon the environment that you--the passive film on Alloy 22 will change as you change its environment, and it is in fact different than what you will typically see for something like a 300 series stainless steel under similar conditions.

NEWMAN: What is the causal connection between the composition as measured by x-ray photoelectron spectroscopy and performance?

FARMER: Well, let me pose a question to you. Why when you add molybdenum to these nickel based alloys, as you increase molybdenum, why do you have a change in the threshold potential. If the alloy elements have no impact on passivity or the stability of the passive file, why does that occur?

NEWMAN: Well, that's a topic which has been intensively debated in the small community of what I call academic corrosion scientists over the last ten years or so. So if you haven't been to those meetings, it would take me too long really to go into it now. I don't want that to sound like a nasty comment, but really that topic has been debated intensively in the last ten years, and there are two schools--

FARMER: But what is the answer?

NEWMAN: The answer is that in certain cases, not in this particular alloy, but for example in the case of 304 versus 316, it's been demonstrated quite conclusively that the whole difference in corrosion performance can be related to the propagation stability of small pit type cavities, and not to some difference in the supposed quality of the outside film. Now, I have not carried out that--

FARMER: But these are not--these films, if you look at them, structurally they're not just chrome oxide.

NEWMAN: They have other things in them, but the--

FARMER: They're mixed films.

NEWMAN: I will just--well, this would be rather an abstruse argument if I was to go into too much detail. But basically, the--

FARMER: What is the composition of the passive film on Alloy-22?

NEWMAN: Well, I don't really care because I look at the problem from the opposite perspective. That is, if I get a certain elevation in properties as a result of adding an alloy element, I examine whether I can explain that elevation in properties, whether it's a breakdown potential, or something like that, exclusively by examining the effects of that alloy element on the dissolution process, the corrosion process that occurs inside the cavity, if I can explain that whole elevation in properties as a result of considering the dissolution in the acid cavity solution, and I don't need to think about what effect that alloying element might have had on the film.

And in the specific case of molybdenum, I believe it's possible to show that irrespective of what differences in composition you might find, that that passive film is no more or less protected than the passive film on even the cheapest stainless steel that you can buy.

FARMER: Well, actually molybdenum oxides are stable at much more pHs than chromium oxide.

NEWMAN: Yes, exactly. That's where it exerts its effect, is in the acid environment of the already developing cavity.

FARMER: The same is true for tungsten.

NEWMAN: Exactly. I wasn't really expecting that to be a super-controversial remark, because actually I think within the--

FARMER: Well, let me ask another thermodynamic based question. If you get into a regime where you would not have stability of chromium oxide but you would have thermodynamic stability of molybdenum and tungsten oxide, would you expect that hypothetical alloy to passivate with molybdenum and tungsten oxide, or would it be immune or would it just spontaneously corrode?

NEWMAN: It certainly wouldn't passivate. It would corrode at a lower rate.

FARMER: Even though it would form an insoluble molybdenum or tungsten oxide?

NEWMAN: Yeah, that's not the same thing as a passive film. That's why it has a lower corrosion rate, is because it forms that stuff inside the pit cavity, or the incipient pit cavity. Actually, I think that particular point is one which I'm happy to leave to sort of the community, if you like, of the longer term, because I don't think it's particularly critical to what we've been discussing.

But I happen to believe that that has been demonstrated.

FARMER: If what you just said is true, and you have a small microscopic pit form in let's say a chromium oxide film, what possible role could the molybdenum or tungsten play in increasing passivity or the ability to repassivate?

NEWMAN: Well, the ability to repassivate is associated with the--it's a coupling between reaction and transport. The process, as you mentioned, I think itself is a kind of autocatalytic process that's catalyzed by the dissolution products of the metal. If the metal dissolves slower because it's got molybdenum and tungsten in it, then you need a much deeper cavity to get the same enhancement of the dissolution products and, therefore, the same catalytic type action on the dissolution.

SAGÜÉS: I would come in at this moment. Maybe I should translate for the rest of the audience, but in case you haven't realized, the presence of about between 10 and 20 per cent molybdenum in these alloys may make quite a bit of a difference, depending on which end it is of those ranges, as to how those alloys perform over long periods of time, and how successful will be the chances that the passive layer will reconstruct itself if it is damaged, for example.

And, again, this underscores a little bit the fact that an extremely important component on the repository scheme depends on understanding what is happening at pretty much often at the atomic level in this system. The understanding is developed up to a point, but it still is limited, and certainly continuing research in this area is important to make sure that we develop the kind of confidence, to use the word, that is needed when we're going into very long-term extrapolations.

I did want to make one point perhaps on something that does not involve very precise mechanistic issues. It's more of an empirical observation. And that is that the kind of alloy that the waste package is made of, the outer two centimeters, the Alloy-22, is an alloy that together with a number of others, was designed primarily for performance in high chloride, low pH environments, places such as refinery environments, and the like.

There is an increasing amount of information, and Joe Farmer presented today some of it, that the immediate environment next to the package surface, because of evaporation of the species involved, may end up being a relatively moderate to high pH environment under certain conditions. And in that case, we may see phenomena that really we're not getting to worry about until maybe the last six months to one year. For example, we may see an enhanced rate of dissolution of Alloy-22 and a potential, at least a little potential, which are not terribly far removed from the expected electropotentials that Dr. Farmer was showing today.

And this may bring up a number of questions that may need to be perhaps resolved in the near term, and I was wondering if Dr. Farmer could comment on that, if he's still around, the question of the peak in anodic dissolution in Alloy-22 at around 400 millivolts when you are in the SCW environment, I believe.

FARMER: Yes, frankly, we don't--we're confident, or reasonably confident that that doesn't correspond to any catastrophic breakdown on the passive film like if you get a pitting potential or something like this. But there's probably some change, you know, an increase in the oxidation state of some metal cation in the oxide film, and we're not sure at this point exactly which cations are changing oxidation state. We're studying that with an x-ray photoelectron spectroscopy and hope to be able to resolve that, because it's important to know. But we haven't answered the question yet.

NEWMAN: You apply an allow, you apply a series of alloys which have one of the elements at a time removed. For example nickel chromium, tungsten, or nickel molybdenum.

SAGÜÉS: That's a very good suggestion.

Okay, it's been suggested to me, and I think that's a very good suggestion, that we should begin to--the last stages of this roundtable discussion, and I would like perhaps to ask each participant to summarize maybe the key conclusions that he or she may have reached in this discussion, and we can do this on the structure or--I like the structure model. That way we can keep--and since Dr. Andrews spoke quite a bit about models and validation to them, he should be the first one to talk, and we'll continue around in this direction, and I'll be the last.

ANDREWS: Okay. Just so I don't use the word in my presentation and talk about multiple lines of evidence that give one confidence that the models are appropriate for their intended use. And I think the more lines of evidence from diverse angles, which includes, you know, analogs, if they are appropriate and available for the different informations. The analogs may not be used in a quantitative sense. They may be only used in confidence building sense, in a qualitative sense. Confidence is added by external reviews of the science, the fundamental underpinnings of the models.

Those external reviews can include expert elicitations. They don't have to. But clearly some of our models which we subjected to expert elicitations for the VA, I think benefitted from those. In fact, that was one of the reasons, not the only one, but one of the reasons for discarding the saturated zone model that was developed for the VA as not representative and not reasonable for the intended purposes, i.e., not valid, if somebody wanted to use the word valid.

Other multiple lines of evidence are multiple indirect or direct observations. I think Bo had a number of them. Joe treats it slightly differently and goes after an issue potentially detrimental to materials performance and tries to get into the lab, into the theoretical basis for that issue, and either determine it's a real issue and incorporated in the model, or discard that as an issue because of data and theoretical basis.

So I think all of those things, the theoretical basis, the direct observations of that process, peer reviews of the individual components by the scientific peers of the people who are grading the models, all combined give confidence. And then when those models are used, the uncertainty in those models which has to be described and summarized within the context of the model can be evaluated, and the significance of that uncertainty to the decisions that are at hand can be evaluated, and allow the decision makers then, based on all of the evidence in front of them, to make a reasoned decision as to how to proceed.

SAGÜÉS: Thank you very much. Dr. Eisenberg?

EISENBERG: I guess one thing I'd like to say that I'm gratified that DOE is using elements of the White Paper strategy that was issued by NRC and SKI. I want to remind everybody that there's two parts of the evaluation of complying with the performance standard. There's the quantified performance of the repository, and there's then also the evidence for confidence in that calculated performance, and those are not necessarily the same thing. They're two distinct items.

I'm not sure, there was some discussion earlier today that you might use the same kind of language, because they both can be described probabilistically, but I'm not sure that the confidence in the models used to project performance is always appropriately discussed in quantitative terms. But qualitative terms might be more appropriate.

With regard to the NRC regulations, I think we expect a reasonable approach. We do not expect the impossible. Part 63, like Part 60, asks for support of the models. It does not ask for validation.

I think there's a need to focus more on extrapolations in space and time, because that's the central issue.

We strongly support the use of multiple lines of evidence to support the models, and I agree with Bob. And finally, just a reminder that reasonable assurance for protecting public health and safety is based not just on the results of the performance assessment, but all the evidence before the Commission, including elements of siting, continuing stewardship of DOE by DOE of the site, and other protective measures.

PARIZEK: I'm interested in just keeping my eyes open all through this process, and the program has to do the same, looking for always some new reason to maybe pursue something that may be an important goal, and that is to make sure we haven't overlooked some critical point.

For instance, that 1000 millimeter flux rate that might be needed to create drips, if that statement is correct, that buys a lot of protection. And if the shape of the tunnel doesn't make much difference and that can be demonstrated, we feel even better that we're not going to have drips.

But then if we go to the test site and we see water leaking off the roof of tunnels and splashing in different places and we say what's wrong with that place. I mean there's a disconnect here somewhere. We want to make sure that we can take and transfer those observations to a place like Yucca Mountain and understand under what conditions we saw water pouring into N Tunnel, G Tunnel, or some other tunnel.

So this is the thing that always works me if something inconsistent has been stated perhaps, and we need to understand the process.

And then the multiple lines of evidence already stated the fact that for the unsaturated zone model, there are many, many different ways in which the model is being looked at, and I think that does add to me confidence that perhaps it's not just the temperature, it's not just the gas, the pneumatic responses, and all of that's consistent with some level of understanding and how that mountain behaves in the unsaturated zone. We need to do the same for the saturated zone.

As far as the metallurgists, they have to do the same for theirs. And then we have to put all this together, and then we'd have a very complicated thing to sort of sort out and say, well, I think at the end, I feel better. But why not allow for the fact that we can change our mind. I think that's a public credibility problem. I think it allows for the fact that perhaps you're going to keep the door open longer than the program originally envisioned.

And there's a lot of good to be said about it, and if people say, well, that's because we don't really trust us, you're never going to take it out, you put it in there and we're not going to trust the program, you have no intention of taking it out, but scientists would say, well, we know we're going to improve our understanding of processes in the future.

We're making progress every day. Our computers are bigger. Our experiments are continuing. And so we always upgrade our science and change our mind, so why can't we convey that to the public, that if you put it underground, the license says maybe that you can take it out, or have to take it out if you find something wrong with it, but the public understands that there is a control over this process and that really it's not just a random decision. You put it there and you have no intention to take it out.

You may be more than happy to take it out after you begin observing the performance of that place, because that's the other part, once you make an engineering decision, you have to kind of monitor its performance to see if your understanding was correct. And if not, you'll make adjustments. And the science and engineering community will make those adjustments, in my opinion.

So I'd hope that we can perhaps do a little bit more with the public perception of how this process might work.

SAGÜÉS: Thank you. Linda Lehman?

LEHMAN: Linda Lehman, Nevada. I guess because of the differing expectations, we should not use the "V" word. But because we do have unique solutions to some of these equations and processes, that we should embark on the confidence building approach, which works to constrain your answers, and as everyone said, through various independent lines of different results or different data bases, which can be compared.

I also think that I should say something about retrieval and contingency plans, which was brought up earlier. Even though we have a retrieval in the regulation and in the law, I don't--I have never really seen a plan for where that would go or what would happen to it. And I know in the real world if we're doing a design for something, we have to have a contingency plan, but we also have to put up some money for that contingency plan. So that's something else might build confidence in the community.

I also think we need to do more confidence building on some of the processes or things, barriers I guess that are the primary barriers, like the waste form or waste package, which are expected to last hundreds of thousands of years, or at least 30,000 years is the latest I've heard. But those kind of time frames are very, very frightening to the public, and I think there has to be a lot of confirmation going on in terms of how long those barriers would last.

APPLEGATE: All right, what have I taken away? We have a failure to communicate. First off, Congress did not intend to be laying out an impossible task. A lot of people wonder what Congress was intending. But the one thing we're certain of is that they were not laying out an impossible task. But it seems to me that validation really does just that, effectively undermining all the calibration, all the testing, all of the work that has been done and has gone into this effort, and which ultimately common sense dictates is all that can be expected, because this is indeed a completely unprecedented undertaking.

I mean, the question that was raised earlier, in that way, it is fundamentally different from, say, building a bridge or what not, because the first several hundred thousand bridges that were built certainly weren't forced to undergo the kind of incredibly rigorous oversight that this project is having to undergo on its first time out.

I agree with the others that to build confidence for the LA, and I'm restating what I stated before, certainly monitoring, thinking of the long-term, looking at contingency, all of these things are very, very valuable. But, again, in terms of a political decision, they're not. That's just sort of the painful reality of it.

So given that fact, and given the fact that you have to accomplish this, how do you build confidence for this political decision? And I think what I really took away was the comments made this morning by Debra Knopman. It comes down to communication, it comes down to understanding how to present all of the work that has been done. And I think that was a very valuable discussion and we're embarked, I'm working a lot on the climate change issue which also deals with models, also deals with people with very different opinions and a seemingly intractable problem.

And one of the things that we're trying to understand is we're doing focus groups with policy makers, trying to understand what their perspective is and what their expectations are with respect to the science. So I think that's quite a valuable undertaking.

So, anyway, that's my two cents.

SAGÜÉS: Thank you. Dr. Tsang?

TSANG: I just have one viewgraph.

SAGÜÉS: By all means.

TSANG: First, I want to make very clear it's a personal view. I do appreciate Yucca Mountain paid for my trip, and also appreciate that you're not giving me a single phone call to say what am I going to say.

But also you did not ask me what I'm going to say, but that is the LBL practice anyway. So my main comment on my experience in INTRAVAL, DECOVALEX, and also I had to write some review reports, review NIREX and Site 94, and I also looked at the Japanese H-12 report, but I don't have the right review about that.

But I will say Site 94 is a very good report one should look at because it discusses lots of the issues.

The next viewgraph, the next part of the one viewgraph is probably not that kind of show, I hope given they will agree. One thing I want to make mention is this contingency plan business. Over 15 years ago, I think, I was in DOE Headquarters. I was asking how about firefight brigade concept, and the answer is no, no, no, don't talk about it. The main reason was that at the beginning of the discussion of nuclear waste disposal, the concept came out is that we want to put nuclear waste away so that nobody after, say, 50 years or 100 years, whatever finite time period, no people need to worry about it. We don't want to burden the future generation.

Scientifically of course I agree with that. There needs to be some kind of monitoring and contingency plan, but we are really going back to the very beginning, the philosophy of the whole thing, so we have a long battle to fight.

The second part I think was covered in the discussion already. The best PA model may not be the same as the field calibrated model. I think we talked about that, so it's very important to have the PA model correct, whatever that means.

Let me just look at these. The PA model result must be given with uncertainty ranges, and the uncertainty is not just parameter value, but also the FEP, the features, events and the processes, and there is a need for an alternative model, and I think I showed the SKI's approach where they look at alternative models and find a discrete fracture, and a simple single fault problem, and even within that, they vary the different conceptual things. And that the uncertainty is different from parameter variability. Those are two different things.

Then in my mind there is a question of how do you bring the state of the knowledge of the scientific community into the PA. That basically I will say is intrinsic limit of model validation. There's nothing you can do beyond that. And then I said it's important to recognize there are three types of experts. One is there is an expert at the Yucca Mountain site. I mean, they've been living, breathing there for the last I don't know how many years, and if you want to know what's going on in the site, I mean, they're the expert.

But it's important to bring the general scientific community expert in and to help with the system so that we are at the forefront of the science. And in the NIREX, as well as SKI, they have a formal system using external experts, not just as a peer review, but also in part of the decision making process in the middle about importance of features, events, about all the impacts, so there is a formal process there, and they document it, so they revise it, everything is traceable and transparent.

And then the other source of expert which is very important to draw from is the nuclear waste expert from other countries, other people's programs. One difficulty about getting expert advice is that in a country, maybe not so much in the United States, but in other countries, almost everybody is working in the waste. They don't have the other experts to draw from. But on the other hand, it would be very useful to draw from experts from Sweden, U.K., and so on, and I note you people from Canada. But I think these people that have been worrying about the nuclear waste program in their own company, they're very good, so they'll be familiar with the philosophy and all that. Now, of course then scientific publications. That is open to everybody, and it's really important.

Then I have some open questions, just three more. How to validate probabilistic model, and that is not so easy. One could look at a range, compare the range. That's one way to do it. There is quite a lot of literature in system engineering, Oren, Sargent, system engineering, there's whole proceedings on simulation, conferences, symposium, where to look at various tests for these kind of things.

I really have difficulty with this one. I don't know whether anybody--how do you validate bounding calculations? Some of the bounding calculations from zero to the sound is probably obvious. But if you want to shrink it and narrow it down, it becomes quite subtle, and that is a hard problem I don't know how to solve. And I'm still pushing that it would be very useful to use multiple independent groups. In the Site 94 report from SKI, they actually used different groups to look at different conceptual models, and each group did the tests and then compared the results. And I think this is one way to try to bring forth science.

So, again, this is a personal view. I don't represent anybody. I'm sure I step on maybe Yucca Mountain and NRC and IES's toes. If you don't know if I step on your toes, you can ask me and I'll tell you.

SAGÜÉS: Thank you very much. Dr. Runnells?

RUNNELLS: I think much of what should be said has been said. From a personal point of view, I'm very favorably impression with what we saw today in terms of modelling efforts and modelling benchmarking, modelling calibration, modelling verification. There's a "V" word, but it wasn't validation. So I thought the presentations were excellent and it shows a great deal of progress.

I sat, though, and I still do sit through these meetings and wonder how much the general public could possibly understand of what goes on here. And in the final analysis, I believe the general public will have the final say. I think that there has not been an adequate, if you like, involvement of the public, or an adequate education of the public so that they can understand to the degree possible the science and the effort and the meaning of things like uncertainty in this program.

I'd take an additional step. I'd say that none of us can understand 10,000 years, none of us. If we think we can understand 10,000 years, we are quite foolish. I think back to what do we know about the time of formation of this country in 1776. How much do we know about what was going on in 1776? That's only 200 years. How much is left for us to view from the time of the Egyptians? Precious little.

We do not understand 10,000 years, and I think we have to recognize that on the front end, to me, that means we recognize that these models are the best tools we have, but that we have to incorporate into the predictions monitoring, appropriate monitoring, and I would argue that we need to talk about reversibility or retrievability, whatever word you want to use, but if something goes wrong, what are we going to do about it. That's what the public I think would like to know.

I'd suggest there's a fourth group of experts, by the way. I would suggest that the public is the fourth group of experts. The public, we as the public, I'll include myself, are expert in how to raise our children, not really, how to raise our dog, how to grow a garden, how to enjoy the out of doors. There is that fourth group of experts that I think this program tends to gloss over. They don't understand perhaps the science, but they understand things that affect their daily lives, and I think we have to pay more attention, the program should pay more attention to them.

I heard mention the other protective measures, other protective measures that might be taken. I'm not sure what that means, and I'm sure the public doesn't know what other protective measures might mean. I think we have to spell those out, whatever they are, in terms of safety to the environment, safety to the public.

I would also submit that this program is not unprecedented. I would submit that the program to take a man to the moon was of equal magnitude and equally unprecedented, but that the difference was leadership. John Kennedy when he set the goal of going to the moon rallied the people behind him. I think those of us of adequate age can remember his speeches and can remember the excitement that the leadership of this country gave to the moon program, totally unprecedented.

Many people would have said it was impossible, you can't do it, and yet with the proper leadership and the proper education of the public, it was accomplished. And I would like to see that kind of leadership again at the very highest levels with respect to this very important and very difficult problem that faces the world of nuclear waste, and I don't see that we have that leadership. I think that's missing. I don't know how we get it. I don't have an answer as to how, but it's missing.

So anyway, enough sermonizing. Those are my thoughts.

KONIKOW: Konikow, USGS. I think I've probably made my position on model validation clear. But I also want to make clear that I do believe in the value and use of models. I certainly didn't mean to imply that I have any criticism of basically the idea of using models to make predictions. I think they are the best tools we have, and they should be used. They should be tested, and they should be viewed with healthy skepticism, and there is a call for letting the public know what we're doing with the model, and we have to understand what the models are doing.

And so--and this is good and it's sometimes hard to do for some of these individual complex models. I mean, the unsaturated zone process, they're very complex and non-linear. So if we think that's hard, wait till you couple all of these multitudes of models into the TSPA system or into the PA model. Just wait till you get them all together. And I don't think anybody in this room is really going to know what's going on in that coupled set of models.

And the idea of a PA or a TSPA is really a good one. In theory, it sounds great, and difficult to argue with it. It's the way to go. But as with many other things, the devil is in the details and I'm perhaps a little biased by having served on the National Academy's WIPP review committee for about seven years while they were going through their PA exercise, and it was great in theory, but there were some real problems with the implementation, with the details, and with the review group like this that meets a couple of days every few months, it's really hard to get into those details. And if you're not looking at those details, well, who is looking at the details other than the people running the PA model.

Some of the problems that we saw, maybe I should just say me, there were some times a disconnect between the scientists on the project who were developing these complex, sophisticated calibrated models that seemed to be representing the processes pretty well, and the abstractions of those models that were incorporated into the actual PA that was making the predictions. Sometimes the PA people weren't talking to the scientists who were developing the original models. This is one of the dangers.

Sometimes it was the way they were doing the sampling procedure for this whole Monte Carlo approach. There are subtle ways that that could introduce bias into the generated risk statistics. There were cases--well, in general what they were doing was independent sampling of all the parameters. Well, if you have two parameters that are highly correlated, then the independent sampling is going to be generating a fair number of infeasible combinations of parameters, and if those are the ones that are generating, let's say, safe cases, what you're doing is stacking the deck. You're affecting the outcome in terms of the risk statistics.

What was being done in some cases was substituting larger variances in parameters for ignorance. You know, one of the things that concerns me about dealing with the natural systems around Yucca Mountain versus dealing with the engineered barriers, is that the range of uncertainty in characterizing the natural geochemical and hydrogeologic properties is really so much larger in terms of the uncertainty in characterizing the engineered characteristics, the engineered barriers characteristics.

And I'm not convinced that we could adequately characterize the mean and the variance and the trends in these properties, or that we could substitute our ignorance of these by just increasing the variance. One of the things is that, you know, for some parameters, instead of representing the heterogeneity, they would just vary the mean value, but keep it uniform for each simulation, for each realization. I would argue that they're not equivalent. They do different things. And that will, in effect, bias the outcome in one way or another.

And so I think that there are--I could go through a whole list of these, but there are a number of subtle problems in the actual implementation of a complex PA in which multiple models are linked together that I caution you to be wary of.

SAGÜÉS: Thank you very much.

ORESKES: Much of what I have to say has been said before, but I'll just try to reiterate a couple of points. It seems to me there's still one issue to be raised that hasn't been mentioned over the stance of DOE towards new information. In the last couple of days, we heard several people say that in the coming months, various tests would be done or various model calibrations or whatever you want to call them would be done that would increase the confidence in the position. And that makes me feel nervous because it seems to me it's putting the cart before the horse, and it raises the question that I think was asked by the Board several times in the last two days. How do you decide whether or not some results ought to increase or decrease your confidence in the situation? What would constitute grounds for decreasing your confidence? What constitutes grounds for rejecting a model? And what are the criteria by which something is determined to be reasonable?

We didn't really ever hear the word unreasonable or acceptable. We never really heard the word unacceptable. So I would just encourage the people involved in this process to think again about that question. And I think that in terms of public confidence, unless one has some sense about what the criteria are by which something is deemed reasonable or unreasonable, then there's this concern that arises that, you know, almost anything could be reasonable if the people decide they want it to be.

So I really raise that as an important issue about the stance of DOE towards the information generating process.

The second point I'd like to make is just to reiterate this issue about the predictive accuracy of calibrated models. A calibrated model can be predictively accurate. There are many, many good examples in the history of science of scientific theories that made extremely accurate predictions, but were later shown to be conceptually flawed.

Several times we've heard the issue about the underlying process, and I think everyone here agrees that we want to understand the underlying process. I don't think there's any disagreement about that desire. But how do we get to that? That's the real question. And the fact that the model may have predictive accuracy is not the answer to how we get to the underlying causal issues.

So I would encourage that issue to stay on the front burner and to hear more talk about the independent evidence for the causal processes that are being invoked in the models.

And then the third point is to reiterate the point that Dr. Runnells made. We are trying to make a decision here in the face of substantial scientific uncertainty, and we could have a really interesting discussion about the space program and the way in which it's similar or different, and I take your point that it was unprecedented in certain ways. But I would argue that the scientific uncertainty is actually greater in this case.

But whether it is or it isn't, it's clear that there is tremendous scientific uncertainty in this process, and then that argues the need for an ongoing learning process, the possibility of preparing for monitoring, modification, retrievability, reversibility, whatever word you like, and it seems to me that as DOE moves towards the final TSPA, that it's really important these uncertainties not be swept under the rug. It's not wrong to be uncertain, but it is wrong to be dishonest about being uncertain. And I think DOE should find more effective means to communicate this uncertainty to the people whose lives are potentially affected by this, because that is what we're really talking about here, and I think it's easy for us as technical experts to gloss over the concerns of the people who live in the state of Nevada and elsewhere. Their concerns may be exaggerated. Their concerns may be irrational by the standards of statistical analysis, but they are real concerns, and I think it's really important for us not to dismiss those concerns, whatever their sources are, and that the DOE should emphasize that this process of learning, monitoring and possibly modification won't end with the site recommendation.

SAGÜÉS: Thank you very much. Dr. Newman?

NEWMAN: I didn't know anything about hydrogeology, or rather I didn't until about a month ago. And the reason I know more now than I did a month ago is not because I've been reading all the documents that I was sent, although of course I did, but because I own a Victorian house with a cellar and I don't walk through puddles of water to get to my wine, and so I decided to have part of it sort of siliconed. And it's remarkable how much you learn about hydrogeology by doing that.

For example, you silicone part of the wall, and then the water starts coming out somewhere else, but I'm sure these things are very obvious to you. Or when the workmen inexplicably disappear for three weeks in the middle of the job, then they have to start again because the whole things comes off the wall.

But it did make me think that perhaps, you know, we're very used--I don't want to sound condescending towards the public, but we're very used to talking--to showing pictures of things, but I'm always much more easily convinced by a physical model. I feel like it's sort of an analog model, if that's the right expression, than any number of pictures of schematic drawings of things, and I just wonder whether the concept of how the water gets into this repository and what the physical processes really are that are involved in it couldn't be explained using a physically realizable model. That's just a random thought.

But going back to corrosion, I think--I just want to reiterate what I said before since I've got jet lag and I can't think of anything new to say, and that is that the most reasonable way to try to guarantee, if that's the right word, a 10,000 year life for these waste containers is to build exclusively, at least to begin with, with what I would call an arrest philosophy. That is, think of all the ways that corrosion could possible start, make it start, and then show that it stops.

And I realize that that's specific to the corrosion issue and can't really be used for the hydrogeology issue, although there is an artist, I've forgotten his name, who wraps things--Christo, that's right. Maybe if you could wrap the top of the mountain just for a few years so that water didn't come in, then, you know, you might be able to carry out a giant experiment which would probably have some merit.

So although it's easy with the little waste container to do that, I don't think perturbation of the natural system should be ruled out either. But then I'm only a corrosive expert.

SAGÜÉS: Yes, indeed. And you mentioned a little bit earlier about the academic corrosion community, and I think that if you put the first two words together, then you get way beyond our field.

NEWMAN: Well, corrosion science is often considered an oxymoron.

SAGÜÉS: Okay, that's very good. We'll we're within two minutes of being on time, so that determines the length of my little contribution.

I really--we have heard a number of very valuable insights. I just wanted my only little comment again in the area of corrosion. We are going to be in need of more basic knowledge on this. There's no question that what causes the passive layer to exist and to remain so, is really not known very well. We don't have--we have a number of very important open questions, and we have one particular issue, and Roger Newman has continued to--in the literature to that and he himself recognizes that this issue still we do not have a fundamental understanding of what causes a given temperature to exist below which processes such as crevice corrosion don't seem to continue.

Now, that concept is critical to a repository design of this type because we're using the concept of a critical temperature and, therefore, susceptibility. And I think that those things are going to have to be known better to instill our confidence in whatever we do, model predictions or otherwise.

But anyway, it's exactly 5:30, and I really would like to thank very much the contributors to the panel. I appreciate very much again all the thoughts that have taken place. And without much more, I'm going to now pass the control of the meeting to Dr. Cohon.

COHON: Thank you, Alberto. Don't anybody move. We're not quite done. Just some brief concluding remarks after a long day, long two days.

I, too, want to thank the members of the roundtable and Alberto for his wonderful job as Chair. It was a very stimulating couple of hours. I got a lot out of it, and I think my colleagues on the Board and others in the room did as well.

Don, maybe one of the presidential candidates will step up and say nuclear waste is the issue I'm going to go public on. Don't hold your breath.

Though we did not engage the audience by design in this, and I'm just another member of the audience, I'm the one who's got the mike so I want to make just one brief remark.

One of the themes that was constant throughout this roundtable was the issue of uncertainty. Unavoidably, this problem is highly uncertain and it's arguable as to whether it's the most scientifically uncertain problem ever attempted. But nevertheless, the uncertainty is very high.

And, furthermore, we've heard some good comments by many people, most recently by Professor Oreskes, about the need to be clear about uncertainty, about the need to communicate it effectively to the public, she mentioned, and that also includes decision makers, political decision makers. And we've heard that comment before, as well as technical decision makers.

It's a wonderful opportunity to say once again, having the expected value of dose is the only decision criterion that does not convey uncertainty. I've raised this before. One answer has been from DOE, well, expected value because it takes into account it's a weighted probability measure, captures uncertainty. That's not true. I mean, that's true, but it does not convey the uncertainty to decision makers.

When I raised it with NRS, the response was oh, well, we're going to present to the commissioners uncertainty also in the full range of performance. But the fact is the decision criteria, the criterion is expected value that's not communicating uncertainty.

One final thing on that note. Somehow the world of TSPA has gotten turned inside out and it's been quite remarkable to watch, and I wasn't really fully aware of it until today. Early on in my time on the Board, there was a wide acknowledgement by the program and especially the people doing the PA, the modelers, that the greatest value of TSPA was to understand uncertainty, to understand a range of possible performance. Now we heard, and the NRC representative said well, I don't think we should be quantitative about uncertainty--about confidence. I'm sorry. That we should be qualitative about it.

Now, the inside out part of this is where they use TSPA to produce a number, the expected value, but we should not be using TSPA to quantify uncertainty. The world has shifted somehow and it doesn't make a great deal of sense to me. There seems to be a large inconsistency.

End of my editorial, and I do get the last word, by the way, at the public meeting. A brief summary of the full two days. A lot has gone on in the last several months for the program, most of it good. We're delighted to see the progress. We're very pleased by the responsiveness of the program to the Board's comments, and we thank you for that. We're delighted by the strong communication links that exist between DOE and the Board and they seem to be working very well, I think for the good of the program.

We heard about the perennial budget problems. They're regrettable and we hope they come out okay. There is no question they will have a significant impact on the program, they must, depending on how they come out, of course, and the time pressures are a constant.

And one other continuing problem is we're going to teach you eventually about the difference between SR and LA, or you're going to teach me that there is no difference.

It was very pleasing to hear about the repository safety strategy and to see the progress that's been made on it, and I think particularly notable was how that strategy and the principal factors that have been identified carry through throughout the rest of the program, and that is what's happening in the field, what's happening at TSPA. There's a sense of togetherness within the program, a sense of coordination that I think is very good, very good for the program, and probably at an all time high.

Thank you again to everybody who made presentations and otherwise participated. My thanks to my colleagues on the Board for their role in helping to chair meetings.

We stand adjourned. Our next public meeting is in January in Las Vegas. We'll see you all there.

Thank you very much, and thanks--I'm sorry--to our consultants and guests in particular who participated in this roundtable. Thank you.

(Whereupon, at 5:30 p.m., the meeting was adjourned.)

WRITTEN COMMENTS BY DR. DONALD L. BAKER

(The following materials submitted by Dr. Donald Baker were included in the Public Comments Section at his request.)