CHAIRMAN HOWELL: We have most of the committee members back.

DR. CALONGE: We had a discussion at the last meeting about very rare conditions and some evidence problems that they specifically posed, and we have been meeting and having discussions by phone with our evidence-based center about how to move forward. So we wanted to kind of go over a couple of these things.

I do want to start out with talking about the potential problems. The first one gets to the SCIDs discussion and even to the last discussion of the assessment of the readiness of public health programs to incorporate new technologies. I don't know how many of you experience, but the cards and letters that came in, after we approved SCIDs, was remarkable to me because people were saying how are we going to do this and where's the technology and we're not doing this now. Someone was saying was this the purview of the committee, and I absolutely think that we ought to think about the process of once we make a vote and we create a letter to the Secretary, I think looking at the assessment of readiness to change, incorporation into existing programs is absolutely something we need to consider.

And the other point is what's come out of SCIDs is an unintended issue that, Rod, at some point we're going to have to

address, and that's that Medicaid won't pay for babies to cross State lines to be transplanted. I think that's a critical issue because not every transplant center in the U.S. has ever transplanted an infant or a SCIDs case. And the fact that we are requiring infants to go to places without that experience because of a glitch in Medicaid reimbursement policy I think goes against what we all understood when we voted yes, which was every child will have access to this if they need it.

So I think understanding the readiness of the public health and the health care delivery system for implementation of point-of-care is critical and actually deserves for SCIDs a specific discussion and I believe a letter. I would suggest a letter to the Secretary saying we need to think about how the federal component of Medicaid policy can promote access to the appropriate care at the appropriate centers for these children.

CHAIRMAN HOWELL: I think we have certainly become acutely aware of the issue in transplantation in SCID and the ability to cross State lines because some infants are likely to be transplanted by centers who have only done transplants for malignancies, and that's a bad thing.

DR. CALONGE: And you say, well, where does that fit in the methods? Well, I think it actually could change the way we think and when we think about voting. And I think we have to kind of figure that out moving forward. So I hope the committee

will consider addressing that.

Then specifically this issue about rare disorders and how evidence gaps are more frequent the rarer the disease becomes.

And then I actually wanted to talk a little bit with the committee about how we could set up our structure regarding our discussion of evidence and science.

So there are three process issues for consideration.

One is standardized process for summarizing and reviewing evidence prior to a vote. So Alex actually did this for us for the last presentation. So for every key question, that slide with the little words on it had that this was the study, this was how direct the evidence was, and this is kind of a summary of what we know around the key questions. But that's a critical table to have, and our request to Jim is that we have that every time because it quickly takes you to the evidence gaps and helps you in decision-making. So I think we can already move forward with that, and I think Jim is going to talk about it as well.

The second I'd like to talk about is executive sessions for discussion of evidence. One of the things that evidence reviews, at least in the other settings in where I work, is that the actual deliberation occurs with the committee itself, free from special interests, advocacy, and politics.

Those are my big phrase. And I think that's an important issue. I think you can maintain transparency and public input, but still have deliberations occur in a closed session. In fact, that actually allows for that. And many federal advisory committees have executive sessions for deliberation. In order to really empower us to have frank discussions about the evidence and the evidence gaps, I think having an executive session in which to do that is important, but it still allows for public comment prior to the vote and transparency of the actions of the committee.

The third thing I'd like to talk about is modeling approaches for very rare disorders. So this is the explicit standardized process for reviewing evidence. One of the things that happened with SCID is that we actually voted on it over two sessions, and we didn't have the entire evidence review or summary when we actually voted on it. And I think in talking to some of the members, making sure that we have this summary -- reviewing the body at each setting will be important.

This is the executive session proposal following relevant public comment, presentation of systematic evidence review, will convene in executive session for deliberations.

And then finally, very rare disorders. There are lingering concerns about hampering decision-making around disorders so rare that our published evidence of benefits and

harms is absent or scarce. One of the suggestions is that we look at modeling the bounds of potential benefits and harms providing us with information useful in considering screening for these rare conditions. What does that mean?

Well, we can use the estimated incidence of disease to determine the upper bound of the number of children who, if identified by screening, could potentially have an improved health outcome. So if there are only 100 cases a year, you can only help 100 children in a year. So that's the incidence issue. It assumes a valid screening test with a known sensitivity, so that makes sure that we can actually detect the children. It assumes a treatment or a management strategy that has the potential to improve health outcomes, and for that we will have to turn to our experts in the area. Do you have a therapy that actually alters the health outcome of the condition? And we can use the estimate of efficacy of treatment and management to estimate the upper bounds of potential benefit. So if someone, for example, feels that out of 100 kids, therapy could help 50 of them, then we know the upper bound of potential benefit.

On the other side, we use the specificity of the screening test to determine the upper bound of potential harms. Those will all be the harms associated with false positive screening tests which require additional testing, potential for

unnecessary treatment, anxiety, and ELSI issues.

The one thing that specificity and false positives doesn’t give you is the concept of incomplete penetrance or over-diagnosis; that is, the detection of illness that, while it’s a true positive, really isn’t going to benefit from therapy. That is, the trajectory, the natural history of the case in that child would be to not need therapy.

But specificity at least will give us the number of false negatives. We can then assess the balance between potential benefits and harms and decision-making, and that evaluation would also assist researchers and public health professionals regarding where we need more evidence in moving forward.

The last discussion about keeping our eye on what happens after we implement something is an interesting issue, especially in the face of rare diseases. I was talking with a friend researcher Nancy yesterday about the concept of thinking about this in terms of phase IV. We could actually implement screening for a rare disease and have a data safety and monitoring committee that looked for adverse actions and then stopping rules, a stopping rule, at which point do we have enough evidence that we know we're making a benefit that we could stop looking. That becomes an ongoing screening issue.

So I want to think about this potential use, and you

could actually calculate it. How many years is it going to take us before we actually know the answer?

So I think trying to think about implementation with phase IV evaluation, both in terms of safety, data safety and monitoring, and stopping rules, at what point does the benefit become statistically significant, are all issues I'd like the committee to consider. We're actually asking whether or not we could have a separate expert working group and present the committee a modeling strategy for very rare diseases and potentially this phase IV implementation approach at a future meeting.

CHAIRMAN HOWELL: Thank you very much, Ned.

Jim, are you going to add?

DR. PERRIN: I guess we would take a couple of questions on Ned's talk at this point.

CHAIRMAN HOWELL: Are there any comments or questions? These, obviously, seem like very sensible solutions to some of the problems we've already discussed today. Mike has a comment.

DR. WATSON: I'm on board completely with finding a way to do it in a controlled environment, data collection activity, whether it be a handful of States that are committed to doing a really population-based pilot.

I'm not so on board with the idea of setting limits because our knowledge of incidence of these rare diseases really

is miserable, the biased ascertainment around it. Every time a condition goes into newborn screening, we find out much that we didn't know before, but we would never know it if we didn't go there. So I think the latter, the phase IV surveillance or some other mechanism of really making sure the data comes in to make sure we're right or actually doing the data collection before we make the recommendation so that we have robust data on which to decide.

DR. CALONGE: So there's this great thing in modeling called sensitivity analysis, and it answers your issue. If our detection rate is 50 percent too low, you can build an assumption in and still come up with an upper bound. So I would say that the issues about incidence and modeling aren't exclusive. You could still do sensitivity modeling and come up with an upper bound if we were half wrong half of the time, if we're missing half of the cases. You do need assumptions and expert opinion to feed that part of the model, but it shouldn't exclude the approach.

CHAIRMAN HOWELL: Further comments?

DR. BOYLE: Quickly. I guess there are two things that were suggested in your talk. One was the idea of an executive committee session, and the other one was a workgroup that would come back to us in terms of thinking about this or actually making proposals. So I guess I would propose that both

of those go forward as a recommendation.

CHAIRMAN HOWELL: Any further comments about that? I think that's a decision we can make without voting on it, frankly. Would you like to vote on it? Would you like to make that as a formal recommendation, Coleen?

DR. BOYLE: Yes, I would.

CHAIRMAN HOWELL: Is there a second to that recommendation?

DR. VOCKLEY: Second.

CHAIRMAN HOWELL: We have many seconds.

Those favoring that recommendation?

(A show of hands.)

CHAIRMAN HOWELL: Those opposing, et cetera?

(No response.)

CHAIRMAN HOWELL: So that is widely supported, and we'll go ahead and do that. And we'll need some suggestions from you and your colleagues about who should be on the working group.

DR. CALONGE: We already have that.

CHAIRMAN HOWELL: If you'll give me the list, I'll look at it and it will be done.

Thank you very much, Ned. That was very brisk.

I'm sorry. Did you abstain?

DR. van DYCK: Yes. I abstained.

CHAIRMAN HOWELL: You abstained. We had one abstention.

DR. LLOYD-PURYEAR: I have a question. Does that mean organization representatives are not in the room? I need clarification.

CHAIRMAN HOWELL: Of what the executive session comprises?

DR. DOUGHERTY: I noticed we're kind of losing the focus on the committee as the committee. The organizational representatives have great input, but I think the committee should discuss first and then take additional comments or vice versa. Take comments and then the committee should work.

CHAIRMAN HOWELL: Is that the general sense of the committee? I see stone faces over on the other side of the room.

DR. DOUGHERTY: It’s nothing personal.

CHAIRMAN HOWELL: Alan, did you have a comment about that?

DR. BOYLE: There must be some definition of what the executive session is.

DR. CALONGE: I think we need to actually look at the FACA rules, Michele.

DR. LLOYD-PURYEAR: I expect there are none, but the committee can close it. I think you guys can define it, but I

will look it up.

DR. CALONGE: So I have a couple of thoughts. My biggest experience is that we actually have, we call them, partner or liaisons in the room for the task force, and while the clinical guide task force does most of the discussion, we certainly value the input of the other folks that are specifically identified as partners. Here, obviously, the other people at the table.

And, Denise, I could see it both ways. I could see actually including the folks at the table in the executive session, as well as the committee members. From my standpoint, that would be valuable.

So I don't know what other people think, but we do need to clarify that.

DR. BOCCHINI: I would agree with that, and I think instead of an executive committee session, you could call it a working group because I think having partner input is very important. Obviously, the vote comes from the committee, but having the input of people with expertise may be very helpful in defining the policy and what comes before the committee as a recommendation.

CHAIRMAN HOWELL: Any further comments?

We probably need to vote on this about what we're going to consider the executive session. Would you like to make

a recommendation?

DR. CALONGE: So I would recommend that the executive session include -- I don't know exactly what you guys are called --

CHAIRMAN HOWELL: Liaison.

DR. CALONGE: -- the liaisons at the table, as well as the committee members.

CHAIRMAN HOWELL: Is there a second for that?

DR. BUCKLEY: I second.

CHAIRMAN HOWELL: Those favoring Ned's recommendation, please raise your hand.

(A show of hands.)

CHAIRMAN HOWELL: Any opposition?

(No response.)

CHAIRMAN HOWELL: Peter, are you voting?

MR. van DYCK: I'm abstaining.

CHAIRMAN HOWELL: Abstaining, okay.

It's unanimous that the executive session will include everybody at the table.

Any other questions? Good.

DR. PERRIN: Thank you. First of all, thanks very much for the opportunity to be with you today and talk about a couple of issues.

Just to follow on Ned's discussion of modeling, we are

very enthusiastic about this set of developments and really want to collaborate very actively here.

I want to take on another set of opportunities to improve the kind of evidence that we bring to you and to talk about ways of providing better grading, and if you think about Coleen's question during the earlier discussion of congenital heart disease, the grading of the evidence on that table we have done some of, but we could do it better and it would help you guys make decisions better I think if we do it. And that's the exactly the discussion we want to get into at this point.

So if you think about our experience together over the last two or three years' time, we've looked at several conditions, and the key questions that you have discussed as a committee do vary by the specific condition. You may remember that when we started working together, we came up with a list of about 8 or 10 or 12 questions that were routine questions to be in any evidence summary. And indeed, every evidence summary that we've provided to the committee and to the bureau has covered all 8 or 10 or 12 of those questions.

On the other hand, in fact, you don't typically discuss all 8 or 10 or 12 of those questions. You've generally focused on two or three or four in each of the conditions under review. A little bit of variation, but it has a great deal to do with test issues, especially test characteristics, what

evidence there is that tests identify, the right populations or not are some of the questions that have been on the table. And the need for some population-based testing data that seems to be relevant has also been a critical question for the group in general.

A second broad issue has been whether there is some value to early identification, and again, we've just been discussing that in the context of critical cyanotic congenital heart disease. Does screening add to clinical assessment?

A third question that we have addressed but not a lot actually in the main conditions we've looked at is does treatment help. Essentially with all the conditions we've looked at, there's pretty good evidence that treatment helps, and that’s not been a major theme of debate in the committee.

And then a last area is the availability of follow-up diagnosis and treatment. The reality here is, as much as we have discussed this, there is usually very little evidence in this particular area. It’s an area where we may want to think together about how to improve the evidence base because it is critical in every condition that we discuss.

Some less critical data, ones that frankly don’t seem to inform your decisions, are the incidence and prevalence of the condition, although as Ned just pointed out well, this is very important in determining bounds of harms and benefits, and

I think as we go into modeling strategies, we'll want to look even more critically at this question.

And then finally, the natural history of the condition itself without treatment has generally not been a critical focus that has led you to make an up or down decision here.

Alex talked a few moments ago about what we believe to be the four critical conditions regarding screening for critical cyanotic congenital heart disease. Does adding pulse oximetry improve the sensitivity? What's the specificity of pulse oximetry? What's the effective early treatment, and how available is follow-up care for test-positive children? And again, this is the same table, just a little different background, that Alex presented to you a few moments ago. And I want to spend a couple minutes on it, again, because it's our first effort in providing you a pretty clean evidence table. We did this relatively recently, i.e., after a discussion a week or so ago about what could be helpful to you.

And it does take those four questions and indicates that for the first two, we actually have some evidence. The evidence varies in some respects, with respect to the quality. The consistency, i.e., do these studies show the same finding, actually isn't that great among these studies. The good news is that there's fairly direct evidence in these two top areas. But when you get down to availability of follow-up care and the

effectiveness of early intervention, we're dealing with quite limited evidence here that doesn’t necessarily get to your questions.

Coleen, we can now but we didn’t go back to the screening table that we provided you and go through that in this way, but that could be a very valuable thing and we will obviously add that to future work.

The main point I want to make here again is it’s doable. This is maybe not the best grading system we want to use, but it’s a doable function. And this shows you in your four critical questions that there are real limitations to the data. They again may not say you shouldn’t make a decision, but there are limitations in the data.

So going back to Ned’s work with the subcommittee here on methods for evaluating conditions, some of the things that you as a group felt to be most important -- and you'll see themes like this consistent in my next couple of slides -- analytic validity, the quality of the data sources and the evidence, the study quality in general, and the adequacy of the evidence or the strength of linkages in the chain of evidence; i.e., does doing something, in this case screening, really improve some kinds of outcomes?

And the quality of data sources tend to be labeled this way. I’m not going to spend much time with this slide

except to say we want to come up with better quantifications, better labeling systems for you as we move forward. And again, in strategies for assessing study quality, there are a number of ways of doing so: the description of the test, the disorder, study design and methods, interventions, adequate description of the basis of the correct answer, avoidance of biases, and appropriateness of analysis, all items that we do look at routinely.

So I want to spend a couple of minutes just talking about one approach, the grade approach. There are a number of other approaches of assessing evidence, but this is one that has a fairly good amount of uptake in a variety of areas, including some real efforts in the last year or two to connect the grade approach to the AHRQ-funded evidence centers' strategies for categorizing data.

And the grade basically gives you this set of grading. So thinking back, Michael, to your questions about what the grading is by the committee now, this is four levels. But this may be one that the committee may want to think about a little bit in its considerations.

So high evidence under the grade approach says basically that further research, whatever it might be, is unlikely to change confidence in the estimate of effect. Moderate evidence is that further research likely will have an

important impact on confidence in the estimate of effect. Low quality is that further research is very likely to have an important impact, and very low is that any estimate of effect is very uncertain.

This just might be a slightly different way of thinking about how you want to grade things and leaves you some opportunity to pick up this question that was present at the end of the congenital heart disease discussion, which is you may have a recommendation, but you want to add some further follow-up that these four notions that Ned talked about that might help with understanding how it improves understanding the estimate of effect.

Now, diagnostic screening and testing. Most of the grade efforts, for example, have really been applied to treatment rather than to diagnostic screening and assessment. Many of the ways of weighing evidence come from whether this particular treatment for this particular kind of coronary artery disease is effective. We're dealing, of course, with a little different problem, which is diagnostic screening and testing here. And the optimal kind of study, if one had one, would be a study that was an RCT of screening versus usually no screening or potentially one screening method compared to another, but these kinds of studies very, very, very rarely exist. It’s extremely unusual for someone to do an RCT of basically

screening versus no screening. And I'll get into the implications of that in a moment.

In general, in thinking about weighing evidence, one wants to think about who the patients are, here who the children are, screened in this context, what the actual intervention is, what the comparison is, and what the outcomes are. This would be the frame against which one would like to look at any evidence for screening.

So the grade approach to labeling quality of evidence basically would label in the context specifically of diagnostic and screening work a randomized trial of the type I mentioned before would be considered high quality evidence. But what's important is that cross-sectional or cohort studies that have good comparisons with appropriate reference standards are also considered high quality evidence. And some of the papers that we do review occasionally rise to that level of quality.

So we essentially never find, as I said, RCTs in this area, but the things that are along this level. The problem in general is that the reference standard may not be a very good reference standard. It's usually an historical control basically, and the biases in that are often substantially high, which does then create some problems in the consideration of these as high quality evidence.

These are some elements I listed below that would

either decrease the grading or increase the grading. I don't think I'll go through them in so much detail except to point out the one in bold there, which is to decrease the grade if there's some or major uncertainty about the directness. What it really means is the evidence that actually doing this screening improves outcomes, not identification. Most of the studies are really going to be indirect rather than direct. So that would, by this system at least, lower the quality of the evidence there. The important inconsistency among studies, as I mentioned, in the congenital heart disease table that we put together -- there is a substantial inconsistency in findings.

So again, using this system as it's currently worked out, really based more on adult screening issues, not on screening for rare diseases, but on this issue, we will have as usual some problems coming up with what this type of group would consider to be high quality evidence.

So the challenge for us here or the Evidence Review Working Group is that almost all studies will be screening versus some published comparisons, no direct comparison, and most evidence would be rated, at least by this scheme, as low or very low.

And I think what we're talking about doing, with your approval and hopefully with using the same working group that you just discussed after Ned's presentation, is to work together

to develop more reliable methods of determining quality. Again, can we come up with a more systematic approach to labeling quality in studies when the quality by the grade system would be typically labeled as low or maybe a bit above low relative to high quality. We don't know if we can do that, but we believe we have the kinds of good minds together who have spent some time on these issues of rare diseases that we ought to be able to come up with a better system of grading.

So in summary, in my presentation, I think what we've tried to do, especially in congenital heart disease -- and I think the last presentation we did too, but instead of having you go through all 8 or 10 or 12 questions, to try to focus on what are the three or four questions that are likely to be most important for your considerations in whether this is something to be added or not to the system. And we will continue to do that. We'd love some feedback as to whether that's useful or not, but we think that's an important part of our work. By Ned's earlier presentation, I think we want to do a much more systematic approach. We want to come up with a strategy for modeling the harms and benefits of screening given what we know about these conditions, and we'd like to develop a better systematic grading and summary of the evidence strategy for you.

Thank you.

CHAIRMAN HOWELL: Thank you very much, Jim.

Are the questions or comments? Mike?

DR. SKEELS: I mentioned this at our last meeting too. It would be very helpful if economic analysis, cost-benefit, cost effectiveness, whatever were included in this. In looking through your slides, I'm back on slide 2, and there is a bullet there that said -- this is key questions affecting our decisions. There is a bullet that says value of early identification screening versus clinical assessment. And maybe that's the right place to plug this in. But in purely practical terms, for those of us who are going to have to go back to our States and sell these ideas, it's very helpful to have some sort of economic analysis. In the absence of that, I can't answer the most basic questions I'm going to get about why should we do this in purely financial terms. So I think it's crucial that we include that kind of thinking and whatever evidence we've got.

DR. PERRIN: That is probably going to come more out of the modeling side of this effort than a knowledge base of clear evidence, but that seems to be right on target and extremely much of what we'd like to do.

Ned?

DR. CALONGE: I would agree. I mean, I was just amazed, the fact that they actually had a cost effectiveness study for pulse ox. So most of the time, you just aren't going to find those. So I think we'll be stuck with trying to put in

inputs with sensitivity analysis about what the test costs, what the treatment costs, and what the costs of delayed treatment are because those are all inputs that we need to look at.

CHAIRMAN HOWELL: We have a quick comment from the audience here.

DR. COTE: A quick comment. My name is Timothy Cote. I'm the Director of the Office of Orphan Products at FDA. We work closely with Kellie and she's our coordinator and official person at the table here, but she comes from CDRH, from the Center for Devices, which is very relevant here.

I come from the therapeutic side, and one of the kinds of evidence that you asked about is do we have a therapy that works. In fact, I'm here today because Sharon Terry said we have to get this together, whereby when FDA spits out a new drug for treatment of a rare disease, there's at least some consideration as to whether or not it has any value for newborn screening.

So I just wanted to introduce myself here for that purpose, and I thought that this question of evidence was a good opportunity to do that. I'll be here later to speak to the medical foods issue. Thanks much.

CHAIRMAN HOWELL: Thank you very much, Dr. Cote.

Is there anything that you presented that we need to vote on? It seems to me that you're going to proceed to do

those things. You're going to be working with Ned with the workgroup that we've already decided to do.

DR. PERRIN: Only if you feel you need to document the charge to this committee.

CHAIRMAN HOWELL: Well, can we have a motion for that effect?

DR. CALONGE: So moved.

CHAIRMAN HOWELL: Second?

DR. DOUGHERTY: Second.

CHAIRMAN HOWELL: Those favoring it, say aye.

(A show of hands.)

CHAIRMAN HOWELL: Any opposition or abstention?

(No response.)

CHAIRMAN HOWELL: It seems to have been passed unanimously.

I have one quick thing. We must go to lunch. We have gotten an extension.

But I might point out that one of our earlier plans does not seem to be very effective, and let me read you from the law governing this committee. We're governed extensively by law.

“Meetings of the advisory committee will be closed only in limited circumstances and in accordance with applicable law. In addition, requests for closed meetings must be approved

by the GSA's Office of General Counsel at least 30 days in advance."

So it seems to me that having executive sessions has probably bit the dust unless you really want to go through this procedure.

DR. LLOYD-PURYEAR: No. We just have to have notice.

DR. CALONGE: We can't do it today, but we can notice it for future meetings.

CHAIRMAN HOWELL: It's complicated, but we can do it, of course. But anyway, that's the law.

Any further discussion?

(No response.)

CHAIRMAN HOWELL: Let's go to lunch and we'll be back in about 40 minutes, 45 minutes.

(Whereupon, at 12:25 p.m., the meeting was recessed, to reconvene at 1:20 p.m., this same day.)