

NATIONAL QUALITY FORUM

Moderator: Lauren Richie
September 20, 2011
1:00 pm CT

Operator: Welcome to the conference. Please note today's call is being recorded.

And at this time, I'd like to turn it over to Lauren Richie. Please go ahead.

Lauren Richie: Okay. Good afternoon everyone. Welcome again to now the third of follow up calls for the Renal Endorsement Maintenance Steering Committee where we're going to look at the dialysis adequacy measures today.

I'm joined here by Karen Pace, here at NQF. And once again we do have another full agenda with six measures to get through over the next couple of hours. So before we get into the measures, I will just briefly turn it over to Dr. Kristine Schonder for any opening remarks.

Kristine Schonder: Thank you. Welcome everybody and thank you for joining us again for this third follow up call here. As Lauren mentioned, we do have a full agenda today to review six measures that - over the course of the next couple of hours.

We will be - the process for those of you who have not been on one of the previous calls. We actually have a very streamlined process where Lauren and Karen will introduce the measures for

us, really kind of highlighting the specific areas of disagreement or concern rather than focusing on the entire measure throughout.

And then as we reach some consensus on the measures we can - we will start to develop a - Lauren will send out a survey at the end to - for everyone to vote on that afterwards in a more formal manner. If time permits then we can talk a little about the related and competing measures and then open it up for comment afterwards - for public comment.

For those - those of you participating in the call if you could please put your phone on mute when you're not speaking that will help to keep down some of the distractions during the call. Thank you and I'll turn it back over to Lauren.

Lauren Richie: Okay, thank you. So at this point, I'll just take a really quick poll of the committee members who are on the call. And I see so far we have Alan Kliger, Kristine Schonder, Michael Somers, Peter Crooks, and I know that Roberta Wages is on the call. Are there any other steering committee members that I did not name that are on the call?

Kristine Schonder: And before we get started, if there are any steering committees are measure developers on the measures today that are somehow in the public line. Operator, how would they signal you that they're in the wrong line?

Operator: If they're in the wrong lines, they could hit the star 1 at this time.

Kristine Schonder: Okay. Okay, so the way we scheduled these calls they're kind of based on the availability of the primary workgroup members. And at the time this call was scheduled we were really only were aware of one person who would not be on the call and that was Dr. McMurray. And I see that we are missing - still missing a couple of the primary workgroup members.

So I just want to - maybe just ask the committee and the co-chairs how do you think we should proceed with missing a couple of our primary workgroup members?

Male: Well, there doesn't seem to be - I mean I suppose one option is to try to reschedule again. I don't know if that would be...

Operator: We just got another, Joe Messana is joining us.

Male: Okay, well that's...

Kristine Schonder: A couple of more committee members.

Male: (Joe and Ally) was on yesterday's call. He reviewed some of these measures. I don't know if he was in the primary workgroup or...

Kristine Schonder: No, not for this group. I see so far we're still missing two, Myra Kleinpeter and Janet Welch.

Male: And as far as you know, Myra was - she did not decline the meeting or she accepted the meeting or...

Kristine Schonder: I did not hear one way or the other and I believe that Janet Welch said she had a conflict after the fact, so. We may just be missing one that may not join.

Lauren Richie: Is Connie Anderson on our group too?

Kristine Schonder: No, no, she is not. So the original workgroup was Myra Kleinpeter, Alan Kilger, Michael Somers, Bobbie Wager, Janet Welch, and Stephen McMurray.

Male: So we have three of those, right?

Kristine Schonder: Three of them.

Male: Well, I'm open to suggestions. Does anybody else in the committee have any suggestions for what we should do at this point?

Male: I think we should press on.

Female: Yes. I just realized I was on mute the whole time. I was looking at the measures and it doesn't look like we have a whole lot of discrepancies through most of these measures. So I would vote for moving on.

Male: I think so too.

Female: Okay.

Female: I think at this time, yes.

Lauren Richie: Okay, well then we'll move on. So just as a quick reminder just remember to maximize your screen in the upper-right hand corner to see the documents a little bit better.

And what we sent out prior to today's call was just a summary of the workgroups preliminary evaluations based on your entries before the in-person meeting. And so it's essentially the same information just in a condensed table format. And as Kristine mentioned, we will provide a brief description of the measures highlighting any areas of issues or concern.

And we do have the measure developers on the call. They'll be available to answer any questions you may have. So I think with that we'll go ahead and get started with the first measure.

Karen Pace: Okay, this is Karen. And I'll follow - go through the measures. We'll go through the measures one-by-one.

Before we get into the individual measures...

Peter Crooks: This is Peter Crooks, I'm sorry. I cut myself by asking - did you guys make a decision on how to proceed? I was going to mute and I hit release.

Male: We decided to press on, Peter.

Peter Crooks: Thank you, yes.

Karen Pace: Okay, so one of the things that we've been doing is looking at generally that the preliminary reviews and based on the measures that were reviewed at the Steering Committee. There doesn't seem to be any issue about dialysis adequacy being a high impact aspect of healthcare. So we've - we suggest that we kind of dispense with not discussing that for each measure. But once to open up the opportunity for the Steering Committee to pull out any particular measure where they would like to discuss impact and then we can do it. But any issue with just picking up with performance gap on these measures and stipulating that they're a high impact area?

Male: I would agree with that.

Female: That's fine.

Karen Pace: Okay.

Male: I'm good with that.

Female: I agree.

Karen Pace: Okay, so we'll begin with measure 0247. And just to recap as Lauren put in the memo, we did discuss three hemodialysis measures at the in-person meeting. And 0249 which was the hemodialysis adequacy measure was voted as suitable for endorsement as well as 0323 which was the physician level measure. So just to put in context.

So these first two measures are also along in the area of hemodialysis. So 0247 is the percentage of all adults greater than or equal to 18 years old. Hemodialysis patients in the samples for analysis's with documented monthly adequacy measures of single pool KT/V are its components in the calendar months.

I think it's pretty clear from that what numerator and denominator are. There's no exclusions. It's a processed measure, no risk adjustment. It's a facility level process measure. And it's a measure of the Centers for Medicare and Medicaid Services.

So if we go on to importance to measure and report in terms of performance gaps, the preliminary reviews rated this either moderate to high. And the developer had submitted some performance data. So I'm going to just stop here at each of the major criteria to see if there are any additional issues about performance gap that anyone wanted to raise.

Male: It remains a performance gap.

Karen Pace: Okay. So in the evidence, basically the preliminary reviews seems to indicate that it met the evidence. One of the things that I think we'll point out is that this would be indirect evidence

because this is a measure of simply doing the assessment monthly which is quite far away from the outcome. And most of the evidence was about the relationship between the dialysis adequacy and the health outcome of mortality.

So it's indirect evidence and it's one of these less than - not very proximal to the desired outcome.

Male: Right. So can I just ask the usual question about that which is that it's necessary but not sufficient measure, right? I mean you have to measure it in order to know that you have adequacy. So this talks about evidence that you're actually measuring.

And we're asked, if I remember correctly to consider each measure on its own and then later to talk about harmonization. So that if we were to find that yes, there's a reason to measure. But it's not proximal enough to the outcome. So that later we want to harmonize with a measure of the actual number - we could do that later. Is that correct?

Karen Pace: Yes. We hadn't really looked at it exactly that way but you're right that we're supposed to look at each measure on its own. And it is, you know, in terms of it is maybe something that you want to look at at the end whether given the fact you have a dialysis adequacy measure whether this type measure would be needed...

Male: That's right.

Karen Pace: ...in that context. So I think that's a reasonable approach.

Peter Crooks: I would - I was also concerned about that. We have three metrics. One that's it's measured, one that a certain method is used, and one that it meets a certain number for both

hemo and peritoneal dialysis. And it seems that it would be possible to come up with a metric that sort of combines them.

In other words to be in the denom- to be in the numerator you have to qualify for these various things. You have to measured - have to be measured the right way. And the number has to hit a certain thing.

And so that it would, it'd be serving the public and the industry better if we could come up with one instead of three.

Male: Peter, that was my point. So that if we look at each of these individually as we've been asked to. Then we have an opportunity later to see if we can harmonize these into a single more useful measure.

Peter Crooks: Right, I agree with the process or understand the process. And that's what we're about. But I'd like to be able to reserve to be able to talk about that later.

Female: Go ahead.

Male: Can I just raise one other quick thing as well? It's the same thing that we've talked about in the face-to-face meetings. And that's the developer pointed out in each of these adequacy measures which is that right now they're describing single pool KT/V which assumes that everyone being measured has the same frequency of treatment.

But these days, we're getting much more variation in frequency. And with that variation a better measure is weekly standard KT/V. so that the developer for this and other measures suggested that we need to - that it's probably appropriate since most people are on 382 look at this on its

own. But then moving forward with this perhaps refine it to a more generalizable measure like weekly KT/V - weekly standard KT/V.

Male: If they did that - is that something they could do without resubmitting it to the committee?

Karen Pace: It maybe something that would need an ad hoc review. And I guess the question would be how quickly CMS could explore that. And we could ask the developer that - I know that they responded to the Steering Committee's recommendation in the prior project. And felt that they didn't have enough time or evidence to do that.

But maybe we can see what they're currently thinking about that is.

Male: Should we ask them now?

Karen Pace: If you'd like.

Alan Kliger: Sure, maybe we can ask - we have Dr. Wolfe and Dr. Messana on the call I think.

Karen Pace: Right.

Male: Could we open those lines?

Karen Pace: They should be on the - operator, if you could open the lines?

Operator: Both those lines are already open.

Joe Messana: So Alan, this is Joe Messana. I'm not sure I understand the question. There were a couple of questions going around. So could someone rephrase the question for me related to standard weekly KT/V?

Alan Kliger: I can rephrase my...

Joe Messana: What's the question - specific question is?

Alan Kliger: Yes. I can rephrase mine. My specific question is this measure uses single pool KT/V as its - no, this, this - let me go beyond this because this is a measure that talks just in general that it's measured not what the number is.

But in each of these measures that have been proposed. Say the metric we're using is the single pool KT/V. The question is would it not be appropriate and thereby encompass many more of the hemodialysis patients to move from this to a weekly standard KT/V instead of single pool?

Joe Messana: So I'll speak for myself and my understanding of where the literature is at right now. Because I don't think I can speak for CMS and if someone from CMS is on the line they may want to handle this. But the difficulty that's - what I'm aware of right now is bad as source for us to do analysis to - to compare standard KT/V certainly from the Medicare data set.

Alan Kliger: All the elements - you have all the elements you need to calculate a weekly standard Kt/V.

Joe Messana: Okay.

Karen Pace: Okay, so then what we'll do is - we'll - I mean this definitely been a theme with the Steering Committee and making this recommendation. So it can be a strong recommendation that - that

CMS move to that. We can ask them, you know, what the possibility would be for them to work on that.

Male: It's probably not something they're ready to do in the short-term. So if - so my question was if two years or a year and a half from they see the light. And or, you know, decide that that's better metric would we - could they just do it. Or would we have to revisit it?

Karen Pace: Yes. They would - what it could be done prior to the scheduled endorsement maintenance. And we would ask them to submit their changes and some analysis to indicate how that changes the performance measure and then a decision would have to be made whether that needed a more thorough review. It's a little bit more substantial than just, you know, fixing a code in a measure.

So it would probably need some kind of ad hoc review process.

Joe Messana: Karen, this is Joe Messana. I'd like to just respond a little bit to Dr. Kliger. Alan, I think you're right. I think in theory we have all the data elements we need even from claims with reported height and weight and frequency of dialysis in single pool Kt/V.

Alan Kliger: Right.

Joe Messana: But I would hesitate to say - to endorse that we could use those data without validating some of the components. The height and weight in particular that I've looked at. I think that we've got a couple of steps there that we need to - we need to pursue to be able to bring a credible measure to you all.

Alan Kliger: Understood, thank you.

Karen Pace: So at a minimum the Steering Committee could recommend that at the time of the next endorsement maintenance you expect to see a measure using standard Kt/V.

Alan Kliger: Okay, thanks.

Karen Pace: Okay. Any other comments about evidence? So and let me just back up to the question that was about looking at this individually and whether - so, you know, there's a couple of ways that you could look at this evidence. And say it's not direct evidence and it's basically expert opinion.

And if that was a strong enough reason not to pass it on evidence the measure would stop there. But I think as we've talked as we go through this the other opportunity will be to look at all these together and make some suggestions regarding harmonization.

Alan Kliger: Right. I mean, again, these - this particular measure deals - it's a frequency measure.

Karen Pace: Right.

Alan Kliger: And I guess in terms of the evidence the thing that was a disconnect for me is that the evidence doesn't deal with whether this is the right frequency or there's evidence that this frequency makes a difference once a month. Whether it be, you know, once every three months, once every two weeks.

There's the evidence doesn't specifically address the frequency issue. That is, it seems to me the point of the measure.

Karen Pace: Okay. All right. Any other comments about evidence?

All right. And then we can move on to scientific acceptability. Here we had basically one difference of opinion about either insufficient or low rating of the evidence. One of the things I'll point is that on reliability, you know, for all their measures, CMS submitted some additional analyses which was in that developer responses to give us some better reliability analysis where they looked at signal to noise analysis and the inter-unit reliability.

So that's some additional information that you all didn't have at the time you did the initial preliminary evaluation. And can...

Male: I guess I'm really not sure their initial reliability when they were doing the month-to-month correlation of the facility score. I guess I'm still not clear what exactly that meant. What they were correlating.

Karen Pace: Yes. And who is that speaking, I'm sorry?

Michael Somers: It's Michael.

Karen Pace: Oh, okay, Michael, thanks.

I think that was the major question. And that's why we had asked CMS and their developer to explain. And that's when they submitted the additional reliability information.

So if you want to ask the developer to explain what they were trying to demonstrate with looking at correlation over time?

Michael Somers: Right.

Karen Pace: Okay.

Michael Somers: I wasn't sure whether they were correlating the specific facility score and how the facility did from one month to the next month. Or, you know, exactly what they were correlating.

Karen Pace: So Joe or Bob?

Robert Wolfe: This is Bob, can you hear me?

Karen Pace: Yes.

Robert Wolfe: So what we were doing there was not trying to validate - look at the reliability of the individual patients' level measure from different sources of data. But to look at the measure at the facility level to see if we would get the same measure under similar - although not identical - circumstances in slightly different circumstances.

And the month-to-month was one way to look at similar circumstances to see if the measure - from serial values - to see if the measure would reliably show the same characteristic for a facility.

And the correlation suggests that, yes; facilities that have good scores in one time period have good scores in the adjacent time period. Now while things may have changed so it's not exactly the same circumstances we are intent with to try and find a situation which was as close as possible.

And this would be like two different reviewers looking at the same patient at nearly the same time to try and evaluate the patient status. Here we're looking at two different measures for the same facility under nearly the same circumstances. And we found good reliability on the score that was reported at the facility level.

That was the intent of that analysis.

Male: Karen, in our table of - I've seen trying to rank reliability testing as strong, moderate, weak, or not applicable. How would you rate this? Is this a moderate rating of reliability?

Karen Pace: Well, I would say the developer's response to our question. I mean it's a bit of a stretch on what our task force has identified as reliability testing at the level of the data elements versus - I mean. So it's, you know, if that...

Male: Might've been weak but then they've strengthened it with their other information they submitted?

Karen Pace: Lauren, could you pull up that table from the developer responses?

So you know, the weakness of that is is that it's at the measured score and not at the data elements. So you could have lots of unreliability at the data element level.

Robert Wolfe: And have consistency from month-to-month so you get the same score.

Karen Pace: Right. And then the other is, you know, in the context of hopefully performance improvement do you - actually a month you don't expect too much difference. But if you were comparing year-to-year, you know, and the expectation was for improvement then what are you actually accomplishing with that?

That was one of the reasons that we had asked them to provide their rational and ask them if they had patient level and facility level data to also provide to you some reliability based on signal to noise which is what Lauren has pulled up for you.

So in this one I believe the entry - what was the entry unit? Reliability for 0247?

Lauren Richie: .94.

Karen Pace: .94 which means that the measures scores able to distinguish well among facilities. In our testing task force rating scale says that it's acceptable for a moderate rating to have reliability demonstrated at either the data element level or the measured score. We would consider this at the measured score.

What they've provided before would be kind of, trying to kind of sit something in that isn't exactly what we're looking for, but.

So I would say that this meets our criteria for moderate rating for reliability testing at the measured score level.

Male: Great, okay.

Karen Pace: Any other comments or difference of opinion about that. You know, in terms of supporting or not supporting either of the methods of testing.

Male: I think this is as good as we're going to get. And the signal to noise is beautiful and I think we should move on.

Karen Pace: Right. Okay. So validity then, again, there was one low rating basically on this one our CMS did looking at the correlation of the quintile scores on this measure which is the frequency measure to standardize mortality ratio.

That's in 2B2. And there was one preliminary reviewer that thought that it did not meet criteria and we can talk about that. It's actually the kind of testing we'd like. But the results, I think - let's

see. Indicated, yes, the results indicated for example core trial 1 with the lowest performance on this frequency measure of 70% the relative risk for mortality was 1.13. Meaning it was 13% higher.

Then core trial - or quintiles, I'm sorry quintiles 2 and 3 were about - were the same at 1.08. And then quintile was 1.09. So it's not a, like, you know, picture perfect progression but does indicate that facilities that have lower performance on this measure have higher risk of their standardized mortality ratio. So I'll stop there and see if anyone wanted to comment on that.

Male: Well, Karen, if that was case of the data we believe. Then we could say this is valid because units that measure on a monthly as opposed to units that don't have a lower mortality rating. And that's a test of validity. Is that the case they're making?

Karen Pace: Yes. Yes. And Bob, you want to add anything?

Robert Wolfe: No that's the main message.

Karen Pace: Okay.

Robert Wolfe: And the - one point is that the lowest quintile, actually, is the one that's really different from the others.

Karen Pace: Right.

Robert Wolfe: So it's not a straight progression all the way up and down. But there is the one extreme value which leads to the most extreme outcome.

For about 20% of the facilities.

Male: Karen, is there different validity testing that would be more effective. Or is this as good as we could ask for or hope for?

Karen Pace: This is actually the kind of validity testing that we're interested in. you know, the results are not necessarily as straight forward as we would like. But at least it does show a difference. But that is kind of the main or at least a - one of the primary questions about, you know, these kinds of process measures is does it tell you anything ultimately about, you know, the quality of care. And certainly correlation with the health outcome it provides some good evidence of validity as an indicator of - you know, that this performance measure's an indicator of quality.

Male: Well, I'm fine with that. Any other questions about...

Male: No, only the note again is I guess, as Bob pointed out. It really is true only for the lowest quintile...

Karen Pace: Right.

Male: ...which is good enough for me but I think we should just note that that's the case.

Karen Pace: Right. Okay. Okay.

And then usability and feasibility do not seem to have concerns but I'll stop there and see if anyone wants to raise any questions.

Okay. And then the preliminary assessment of criteria being met to be suitable for endorsement. Again, there was one disagreement there. And so we can - it may have been related to something we already talked about but - oh, I'm sorry. I'm on the wrong. There was actually three yes and two no in this preliminary vote. So we should talk about that a little bit. And see based on our, you know, what the issues are and whether - what's your thinking now is.

Male: Well are there the two people with no's on Janet or Constance?

Karen Pace: Right. Is that who?

Lauren Richie: Yes. And I pulled her off during the call...

Karen Pace: Right. Okay. Okay.

Male: Neither of the no's is on the call?

Male: Yes, neither of no's is on the call.

Karen Pace: Okay.

Male: Makes it easy for us guys.

Karen Pace: Right.

Male: Okay.

Male: I wasn't sure how they - I thought I'd look back to see and I didn't see the - they passed on all the other sub-measures I thought. I don't know how they came up with the no.

Male: Well, except Constance had low on feasibility I guess.

Karen Pace: Actually go back. I think there was one no on scientific acceptability.

Lauren Richie: That was Janet.

Male: That was Janet, yes, so those were the two, yes.

Karen Pace: Okay. All right but nothing else that this group wants to bring up? Okay. All right.

So we can move on to 0248. And this is a CMS measure. And this is about the method of measurement of the delivered hemodialysis dose. And the description is it's the percentage of all adult hemodialysis patients in the sample for analyses for whom the delivered hemodialysis dose was calculated using UKM or Daugirdas II during the study period. And for whom the frequency of hemodialysis per week is specified.

There are no exclusions and no risk adjustment. It's not applicable for this measure. So we can move into on talking about performance gap and we had agreement that the measure met at either a moderator or high rating. So I'll just see if there are any additional comments or questions about that.

Male: Let me just make one general comment. John Daugirdas actually is just putting together and putting an abstract to use the Daugirdas II for different frequencies of dialysis as well. He has some data that he's showing that's showing you can use the Daugirdas II as long as you can note the pre-dialysis interval - the interval between dialyses.

And you can use the same basic calculation for calculating standard Kt/V. so that it's a method that's adaptable not just for this but in the future. If we move toward a more universal measure like standard Kt/V we can continue to use it.

Karen Pace: Okay. Any additional comments or any comments about performance gap.

Male: I mean since it's virtually the only - it's the only method anyone is using. I'm not sure what we could say about a performance gap.

Karen Pace: Right. So that was noted. And we can look at what CMS submitted.

Male: Although there's a comment on the write up that I'm looking at. It says performance gap analysis performed in more than 70%. Half the facilities use Daugirdas II in more than 70% of patients.

I guess that doesn't tell us what the gap is. It says a lot are using it but it doesn't measure the gap per see.

Male: I don't know what else people are using.

Male: Yes, you know, it wouldn't be wrong to not pass this if there's no performance gap anymore.

Male: Again, in my own sense is I like your idea, Peter, of these are all reasonable things that continue to make sense. And if we can harmonize them all after we pass them that makes a whole lot of sense to me.

Karen Pace: Okay. And, Peter, I know you were looking in the measures submission form. Because they didn't give the facility distribution of scores. So I think in the - let's look, probably in their validity testing for the quintiles, it might be there. Let me just look.

So in their validity testing in 2B2.3 they have quintiles in performance -- quintile 1, 44%; quintile 2 44% to 63%; quintile 3 63% to 69%; quintile 4 69% to 76% and quintile 5 76% to 100%. But I guess the question is - if they're not getting that method does that mean they're just not getting measured perhaps?

Male: Yes. That's what I wonder as well.

Male: Well, these results that are being shown on the screen right now. These are of adequacy, right?
Not of what method was used, or am I wrong about that.

Karen Pace: No. This is supposed to be performance on this particular measure.

Male: Yes. But your point is the right point. That is it's either they performed this or maybe they did nothing. So it doesn't necessarily reflect on the measure it's just, you know, binary. Either they did it or they didn't perhaps.

Karen Pace: Right. So Joe or Bob, could you confirm that?

Joe Messana: I can't where I'm at. I'm away from the office. But if...

Karen Pace: But I would think since the denominator has no exclusion and the numerator is if they had measurement using this. I mean so the option would be the ones that aren't meeting the numerator either weren't measured at all or used something that wasn't acceptable.

Joe Messana: Right. That makes sense to me but I'm going straight from memory right now. And rather than tell you something that I remember but is inaccurate I'd rather get the correct answer to you.

Karen Pace: Okay.

Male: You know, that really brings into question how the measure's specified. Because if this is supposed to be the percentage of patients who were measured, who were measured in a certain way.
That's, yes, you know, there's two sides to this.

But to back to the performance gap issue - I'm - I feel pretty strongly that if there's no performance gap we're not doing anybody any service to pass this. And, yes, in harmonization we might ask them to specify in their harmonized and final metric that this method be used. But it's not proper to pass it if there's not performance gap.

So maybe we should ask the submitters to clarify their data on is there really a performance gap or not?

Karen Pace: Right. But I think that - and I assume Joe or Bob could confirm that. What they gave us for quintiles and performance would be performance on this measure, right?

Joe Messina: I'm 99% sure that that's the case. I just want to check with some of our analysts before I - before I confirm that. But I believe that those quintiles that you're looking at or that you've looked at under the validity were quintiles where UKM or Daugirdas II was used.

Whether there's specificity in the data about other formulary or approaches that were used or missing data I can't comment on at this point.

Male: Yes. But remember that the numerator statement really has two requirements. One that a adequacy number be calculated and the second is that it be calculated using this formula. There are really two parts of it.

Male: Right. And I suspect that it's the absence of the Kt/V but I'd like to confirm that basically.

Karen Pace: Okay.

Male: But then - then there's a flaw. You know, the metric is intended to encourage - to say that the national standard that when it's measured it should be done by this method. Then there's a flaw in the metric if it's not giving us that. You see my point, does that make sense?

Male: Yes.

Karen Pace: Yes. Let me look at the denominator - yes, in center, it just says number of patients receiving in center hemodialysis or home hemodialysis, okay. So all right, so we'll have that discussion or we'll note that. And ask for a clarification.

Peter Crooks: Even if they don't change the metric if they could show us that there's a sizable proportion doing some other method.

Male: Yes. But you know that won't be the case, Peter.

Peter Crooks: Well, I guess I have to decide - at least for my voting how to...

Karen Pace: Right. So I mean, so I think that brings up the question that you all started with is why do you need three measures.

Male: Right.

Karen Pace: If you can combine the - are they testing with how they're testing to the results is...

Male: That's right, exactly right.

Karen Pace: Okay. All right. Okay. So and we have the same discussion about evidence. It's indirect evidence. Any other comments about that?

Again, we have the same issue about the reliability and in this case the reliability, again, the developer submitted that additional reliability testing for the measures score level. And I see that there was one initial low rating on validity.

And I think the quintiles of performances associated with standard mortality ratio are a little less straight forward even then the previous one. So let me just.

Okay, the validity on this, again, you know, gives the quintiles of performance on this measure which is the method. And then correlated with - or associated with standard mortality ratio. And essentially there was no difference really. Because the relative risk went from 1.06 to 1.12 to 1.07 to 1.05 but they all included one which means they were not statistically significant.

Male: Did they all include one?

Karen Pace: I...

Male: I thought they were above one.

Karen Pace: Let me double check - oh, I'm sorry.

Male: They were all over - I thought they were all okay.

Karen Pace: No. Right. No the first...

Male: ((inaudible))

Karen Pace: ...the first one, the first quintile was 1.06 and it's - you're right. It went from 1.02 to 1.09. So the first quintile, there was a small difference. There was a 6% increase in relative.

Male: No, this is a nice - this is a very nice demonstration actually.

Karen Pace: Okay.

Robert Wolfe: And this is Bob Wolfe, I wanted to say that what we have generally presented the validity relative to another outcome. And we have not stated what I think is clear to everybody. That these have a face validity. That these are measuring what they intended to measure.

So that's...

Karen Pace: And I'm - yes, and I need to correct myself. None of these actually included one in the range. So that was my mistake on this particular measure. So they all had increased relative risk, again, above...

Peter Crooks: But I have to point out that because the metric isn't really measuring, you know, that they're using the proper method. This validity test isn't valid. This shows some correlation between, again, whether modeling is done but it doesn't show us that the Daugirdas II out performs other models. You know, we shouldn't read it that way. I don't...

Male: ((inaudible))

Peter Crooks: ...this tests validity.

Male: No, Peter, but it does. You have just put this to the measure as it's written which is that a - adequacy measure be done. And that it be done using the Daugirdas II. The way this is written and it includes both of those pieces. This does test the validity of that.

We believe that's it likely that it's the validity - that people get it measured end up having better mortality compared with the people that don't get it measure. But I think we have to pass on it to talk about it as this is constructed which is a two part numerator.

Peter Crooks: I get what you're saying, you know, but this doesn't add anything really to the last metric then. You know, they're both basically saying that if you measure you have a better outcome.

Male: Well, that's why, again, I'm - my feeling is that your idea is right. That we should if it - if it holds together and makes sense that we should endorse it and then request that these be harmonized into a single measure.

Karen Pace: I think Peter's raising the question, I guess, of whether this measure should move forward, is I guess what you're saying. Right Peter?

Peter Crooks: Well, you know, I'm prepared to deal with the metric as it is as opposed to what it should be, but...

Karen Pace: Right.

Peter Crooks: ...the way it's specified doesn't really get to, you know, saying that we have a national standard that the Daugirdas II should be using. We're going to measure you on them. It doesn't do that. But we will pass it, you know, we will evaluate it as specified.

Karen Pace: Well, that - how it's specified is the validity question. It's not - and maybe this validity testing would have looked more dramatic if it was more clarified but we don't - as Alan said - this is based on the measure as it's specified.

Peter Crooks: But I understand that.

Karen Pace: Okay. Any other comments about scientific acceptability, Lauren, I'm giving you the move on sign.

Did we - and any comments about usability and feasibility.

Male: Everybody said, hi, it looks like, so.

Karen Pace: Okay. And then...

Peter Crooks: I am medium, yes.

Karen Pace: Yes. And then preliminary assessment of suitability there was one no in the preliminary but that was Janet, so. Any other comments or questions? Okay.

All right then we'll move on to 0318. And this is - we're going to start in this - we have four peritoneal dialysis measures. We'll start with two that are the actual intermediate clinical outcome about adequacy.

This first one, 0318 is CMS facility level measure. And it's a percentage of all adult patient peritoneal dialysis patients whose delivered peritoneal dialysis dose with a weekly Kt/V urea of at least 1.7. And in parens it says dialectic plus residual during the four months study period. The

denominator is all adult peritoneal dialysis patient who have been on dialysis for at least 90 days other than that no exclusions. And we can get into the measure now.

So we'll pick up with performance gap. And there seemed to be agreement that it was rated either high or moderate on performance gap. Any issues? Let's see - if there was any other question.

Okay, all right, well then let's go on to evidence.

Alan Kliger: So can I start raising an issue with the evidence, it's just interesting and...

Karen Pace: Yes.

Alan Kliger: ...in several places we've talked about this before. For reasons that are not totally convincing, in peritoneal dialysis when we measure adequacy we include endogenous or residual renal function as is specified in this measure. While for all the hemodialysis measures we've not done that. We've only measured dialysence by the machine and we've excluded endogenous kidney function.

And there was - there has been some discussion about that in the literature. But I think we need to point out that there is apparent - there is that discrepancy. And I for one am unconvinced that there should be a substantial difference in how you look at that for patients on hemo versus those on peritoneal dialysis.

I don't know whether the developers would care to say anything about why that is. They develop both measures.

Joe Messina: Alan, this is Joe Messina. I wouldn't care to but - as you know, you know, these were from 2007. These were based primarily on the KDOQI guidelines and other international

guidelines. And as you know the PD has traditionally included that. It's a weak answer but - but I hear what you're saying. But they were modeled based on the national guidelines that were available.

Peter Crooks: But didn't we use to believe and maybe it's still the case that many hemo patients lose their residual hemo function so quickly that it wasn't worth the time and effort to try to measure residual renal function?

Alan Kliger: Peter, people in the sciences have really questioned the wisdom of that decision. It turns out that there's a big spread of endogenous kidney function or residual renal function in hemodialysis patients. Some lose it very quickly and others don't these days with a more biocompatible membranes, etcetera.

So that the rational of having that split was never very strong to begin with but I think it really begs the question now. I guess I only raise that not to put an impediment in the discussions but only that perpetuating in the year 2011 and up the next two years if we endorse this, that, you know, that difference - basic difference in measurements between the two techniques really doesn't make a whole lot of sense to me.

Joe Messana: Alan, this is Joe Messana. I would only comment - although I agree with you - the hard evidence is lacking that clearly defines the difference between - between hemodialysis and peritoneal dialysis effects or the spread of the residual renal function in prevalent renal patients.

There is the observational literature looking at hemo versus peritoneal dialysis and survival indirectly questions what differences or whether there are differences based on differently mortality in early versus later periods after initiation of dialysis which some people have - have attributed potentially to residual kidney function. But it's indirect and very soft.

Alan Kliger: We shouldn't get stuck on this. And I wouldn't make this the reason not to endorse the measure. But we have to point out the inconsistency and perhaps a message back to developers that in the next round they consider having a more a uniform of doing this.

Peter Crooks: Good.

Karen Pace: Okay. Anything else about evidence. Okay.

Peter Crooks: Well, did they resub- they resubmitted more data on the viability as in the...

Karen Pace: I believe so. And - 0318, did they provide - or there. Is it on that table? We're looking. Yes, it's there 0318, that's the measure we're on, correct.

Oh, the entry unit reliability was 0.57. So lower than what we were seeing for the assessment measures - you see it? No, on that table you were on. Let's see that table. 0318, you see it?
Okay.

So I don't know if Joe or Bob want to comment of what you think's going on with the difference .57?

Peter Crooks: The P value's still very low I don't know why that is if...

Karen Pace: Well, if you have a large enough...

Joe Messana: Large sample.

Karen Pace: Yes.

Peter Crooks: Yes, okay.

Alan Kliger: Also I guess and, Bob, I'd be interested in your thought about this to help me think about the inter-unit reliability number. Hemodialysis has lots of patients in each facility while the PD numbers are very small per facility. Would that make a difference here?

Robert Wolfe: I believe that's precisely it. This is a sample size per facility phenomenon I believe is the major thing that's going on here. But we still do have the ability to identify differences between facilities.

Alan Kliger: Yes.

Karen Pace: And it looks like - I think this was in all of your submissions for peritoneal dialysis information. You say that the mean number of patients per facility was 84.

Robert Wolfe: That's hemo. That's not peritoneal.

Karen Pace: And they're - so I guess that's a question that I had when I was reading this because when we ask for the data. Where the data came from on the performance gaps, they say 34 - 3436 facilities and a total of 293,694 ((inaudible)).

Robert Wolfe: No, that's all hemo and PD together.

Karen Pace: So do - so we don't really know what the numbers have been for any of this PD data that - for any of these measures. Or we have gotten to the rest of the measures but I guess we need to get that clarified.

Do you know what...

Robert Wolfe: We will get that information. I believe it's about 9% of that number is the latest thing I recall.

Joe Messina: Six, it's about 6% total, yes.

Robert Wolfe: Okay.

Karen Pace: Okay. So that - that certainly will affect the inter-unit reliability, you're right.

Alan Kliger: That's what I thought it likely is.

Karen Pace: And then on validity which I think is also related to sample size now that we've talked about it. But if we look at - again, their analysis in 2B2. And...

Alan Kliger: That includes one I mean there's not...

Karen Pace: Pardon me? Go ahead.

Alan Kliger: They all include one.

Karen Pace: Yes. This is the one. I'm sorry I misspoke earlier. This is the one where they - all of the ranges for the relative risk include one. So and you see the spread quintile 1 is 0%, 2 is 1% to 3% - so it's...

Alan Kliger: Well, if they really use just PD - which is what I hope that they did - rather than the whole sample. I'm just certain - again, Bob you can tell me. But I'm certain...

Robert Wolfe: That is correct.

Alan Kliger: ...that is because there is such few patients in each facility that you have very wide countenance in the rules here.

Robert Wolfe: That's right.

Joe Messana: Alan, this is Joe Messana. I was asked to review these data of prior to the submission of additional validation data. And we were only looking at the PD patients here. And that limited our abilities to identify statistically significant differences.

Alan Kliger: Yes.

Karen Pace: All right. Any other questions or comments about reliability or validity for the PD adequacy measure, the facility level measure.

Peter Crooks: Just to help me interrupt the last discussion then, are we concluding that reliability is meeting standard and validity also?

Alan Kliger: I wasn't sure. I mean I originally had said it didn't based on that testing because I thought we supposed to access it based on the data you're provided. I mean everyone else who originally reviewed had polar opposite opinions, I mean.

Karen Pace: Right, well, I think you were - you're correct when you look at the data submitted and the inter-unit reliability of .54 power of 57 not that great. But I think the other thing is to look at it in the context of what Alan was pointing out of the small sample sizes. And what we can really expect with that.

So it is an interesting question of, you know, how to - and I don't think there's any right or wrong answer here. And I leave that open for the rest of the committee to give your thoughts about given these numbers, the quantitative analysis which is the kind of analysis we want. We also would prefer that analysis clearly show reliability and validity. But I think you also have to take into consideration the context and the sample size, so.

Peter Crooks: But doesn't it sort of imply that if I'm a small PD unit and I get this metric measured and I get a value back. I will have, you know, they say, well, you're compliant. You're at 67% but it could be anywhere from 20 to 80. You know, is that sort mean that the confidence interval around any measurement is going to be quite wide.

Alan Kliger: Well, I guess my opinion was that I don't want to get schnookered by the small numbers. It means that we're unable to do the kind of validity testing you'd really want to do. But if - we believe that there is a relationship between individual patients measured adequacy and outcomes independent of this analysis.

I guess I went outside this analysis. I felt that it surely was a valid, was a valid measure. That's the way I thought about it.

Robert Wolfe: Could I speak - add something as well? This is Bob Wolfe.

Peter Crooks: Yes, Bob.

Robert Wolfe: The small sample size means that the confidence intervals will be wide, so Peter, you're exactly right. But in terms of identifying facilities that have a persistent problem. To the extent that a facility has a persistent problem then sooner or later that their data will be bad enough that it will - that it will flag.

And with repeated measuring we will at least be able to identify persistent problems with repeated measuring. We will have a substantial false negative rate that is we will be missing facilities that might have problems. But with repeated measuring we will be able to identify them. So there is value to this even though it is far from a perfect measure.

Peter Crooks: Okay, thank you Bob.

Karen Pace: Okay, any other questions or comments about 0318. It looks like under usability and feasibility the preliminary values were either high or moderate. And no big concerns but anything that anyone wants to raise?

Okay. And is it suitable - there was one no originally and that was Michael?

Michael Somers: That was me for the reason we discussed. I think that, you know, I didn't understand that I could bring in some of the other points that had been brought into this discussion making my conclusions.

Karen Pace: Well, I appreciate you kind of doing it by the numbers initially because it helps, you know, us really air these kinds of tradeoffs that need to be considered. So it's probably valuable to have that for us to look at.

Okay. All right. So let's go to 0321 which is the clinician or physician level measure for peritoneal dialysis adequacy. And this is the percentage of patients age 18 and older with a diagnosis of ESRD. Receiving peritoneal dialysis who have a total Kt/V equal to 1.7 per week measured once every four months.

Peter Crooks: Just for clarity, you know, that I presume - I didn't go back and look at the original submission but that should be equal to or greater than.

Karen Pace: Okay. Right, right, right, thank you.

Peter Crooks: I presume that's the way they submitted it. If they did it this way, then we should send it back to them for editing and rethinking. So anyway, go ahead.

Karen Pace: Okay. Good point. Yes. Okay.

So let's see, exclusions, they have their kind of general exclusions of documentation of a medical reason for a patient not having the Kt/V greater than or equal to 1.7 per week.

And then they give an example such as patient has residual kidney function or other medical reasons.

And we can look at the - when we get to reliability we can look at those specifications in detail if that's something that you want to look at.

Alan Kliger: Well, I mean it has to be. Because either they define it the way we talked about it in the last measure which includes endogenous kidney function. I'm not sure because it doesn't specify it here.

Or if it doesn't include it then you'd need a much more precise definition of who you're not, you know, who wouldn't have it measured. What would the endogenous kidney function be that would make it permissible to exclude patients from measuring Kt/V?

Karen Pace: Okay. So this would be, you know, depending on, you know, the outcome of your assessment of this measure that would be a definitely a harmonization issue as well. But could be explored as well but I hear what you're - your comment was. It either needs to be part of the measure as it was in the previous measure or more clearly specified what the exclusion would be.

Alan Kliger: You know, I have a primary problem, you know, even going on from here I must say. I was one of the reviewers but I may be - it feels a little different to me now. That unless that's clarified it's hard to make a judgment at all. It's hard to go on.

Because the definition is so imprecise that it's hard to know how to interpret it. Either it does, you know, either the numerator statement does or does not included endogenous kidney function.

Karen Pace: Okay. So...

Peter Crooks: Can we open up the original submission or the more the complete submission and look at that - their specifications?

Karen Pace: Yes. So we'll go to the specifications for the numerator. And I believe we have someone on the line from the measure steward.

(Deidre): Yes, we're - this is (Deidre) and I have (Katherine) also on the line for the AMA.

Karen Pace: Okay. So Lauren has pulled that up. And let's see - we'll read what you. Numerator statement was repeated and again it should be greater than or equal to - I think that's the problem with using symbols in our system. But no details except, yes, the EHR specifications which we haven't looked at. I don't know if any of the reviewers got into the details of the EHR specifications.

Alan Kliger: Can we ask the developer if this numerator specifies it includes or does not include endogenous kidney function.

(Deidre): This is the AMA. We have included residual renal function as an exception.

Alan Kliger: So that the numerator statement does not include endogenous kidney function. Is that correct?

(Katherine): Just to clarify that would only be a valid exception if the numerator's not met. So you would first actually look to see whether or not the Kt/V is greater than or equal to 1.7. And if it is not met that could be an appropriate reason as to why it's not met.

Alan Kliger: The calculation of Kt/V is inclusive or not inclusive of endogenous kidney function. That's my question.

(Katherine): I would say it is inclusive because we have not specifically defined it to exclude it.

Peter Crooks: But would - it's inconsistent to say it's included and then you're also putting it as an exception. So I take it to mean that this hasn't been thought through adequately and that total Kt/V does not imply residual renal function.

(Katherine): The way the metric is laid out to be calculated is consistent with the PCPI methodology on how to calculate exceptions. We'd be happy to provide additional documentation on that. But we did not specifically exclude that as a reason, you know, prior to the calculation of the measure.

Alan Kliger: Yes. And again, that would make sense to me but the clarity is that if the numerator includes both dialysence and residual renal function in the calculation of Kt/V. If that is the case as in the facility level measure that we talked about before.

If that is the case and in the AMA measure you then go further to say if it's not measured a reason for excluding it is some degree of residual kidney function. If that's the way you've laid it out then is there any specification of how much residual kidney function would allow you to not measure the Kt/V?

(Katherine): So we have not provided that level of specificity in the specifications. We would be happy to consult with our clinical experts to provide that guidance.

Alan Kliger: Okay. I just think that's a problem in defining the measure. But we can move on. But I surely would assure that we have that specification before we would consider endorsing it.

Peter Crooks: In interest or harm is if it did, was endorsed, I think in the harmonization effort it's quite clear we would ask them to specify in the numerator that it includes residual kidney function just as CMS measures does.

Karen Pace: And the PCPI people the way this is currently specified though it would be each physician deciding what level of kidney function would qualify as an exception?

Alan Kliger: That's my problem.

Karen Pace: Okay, all right. Okay. Well, we'll get - we'll...

Male: ((inaudible)) other medical reasons...

Karen Pace: Right.

Male: That's just ((inaudible)).

Peter Crooks: Yes, that's also very nebulous too. I have - and that isn't present in the facility level metric. And that's an out for - you know, every patient has multiple medical conditions in our population.

(Deidra): AMA we got disconnected there we apologize for that.

Karen Pace: Oh, okay.

Alan Kliger: So we're saying if you didn't hear it - that the specifications as non-specific as they are for the exclusions makes this a hard measure to evaluate. Because it looks like an individual physician can exclude from measurement patients for whatever their personal judgment is about how much endogenous kidney function or any other medical condition without any specifics to that.

Okay.

Peter Crooks: So any physician that's paying attention could score 100% every four months or however often it's done.

Karen Pace: So how do recommend going forward? Do you want to get a response or do you want to talk about any of the other data. I guess the question would be that the data on performance and the testing would be on the measure as specified but we can move. Should we go ahead and at least look at that?

What's - how do you want to proceed, Steering Committee?

Alan Kliger: Well, Peter, you're the chair, what do you think?

Peter Crooks: I'm sort of thinking that if we go through and reject the measure for this or other causes then we don't have a chance to harmonize. And whereas if they came back in a very quick turnaround time with the numerator harmonized with the CMS one. And say this is basically the same metric but this is at the clinician level and other levels outside of the facility. That it would be more attainable to the committee.

Alan Kliger: So if I may, not only that it does need clarifying that numerator. But what I'm hearing from the developers is that their understanding is that it is indeed harmonized already the numerator is. The numerator includes endogenous function they believe. But I'm more concerned about the exclusions that are non-specific and make it very hard to consider this measure.

Peter Crooks: But both, how can you have endogenous renal function in the total Kt/V and then also use it for an exclusions it's...

Alan Kliger: Well, the way they're saying it is that so those people who don't have it measured a reason for not having it measured. And the measure includes endogenous function but if it's not measured a reasonable exclusion is if there is some level of endogenous kidney function.

Peter Crooks: Right which I don't buy as reasonable.

(Deidra): This is the metric developer. If we could just add one thing, it's not whether or not the value is measured. But it's whether or not it is above that range of 1.7.

Alan Kliger: Right. Correct.

Peter Crooks: Right.

Alan Kliger: We understand that.

Karen Pace: So just like - does total - so total Kt/V does not necessarily imply that it includes residual, is that?

Alan Kliger: That's what I'm asking I guess. That was my question and I hear from the developers a probably yes to that - that it does.

Peter Crooks: But Alan, if they're saying that - okay, their total Kt/V was measured. But it was less than 1.7, say it's 1.4. That includes residual renal function then how can you exclude it because they have residual renal function. It's already included in the number. You can't have it both ways.

Alan Kliger: Yes. I agree. I guess, Peter, what I'm saying is because they're not specific about what level of endogenous kidney function - they could be saying that there's a half of a cc per minute of endogenous kidney function which would be a Kt/V of, you know, 1.2. And that that would be a reason for excluding it. Well, I think that's wrong. You follow what I'm saying?

So it is consistent the measure that they present in their arguments consistent it's just that there's not enough specificity in that exclusion to make any sense.

Peter Crooks: Yes. And it causes to question the whole validity or, you know, the - I guess the validity of this as a metric especially if it's endorsed by NQF.

So so, Karen, should we - is it reasonable to ask them to clarify the language surrounding these particular issues and with a quick turnaround time or...

Alan Kliger: Because they're the ((inaudible)) they can't - they're not going to be able to turn around quickly because they're going to need to specify what level of endogenous kidney function would be an exclusion, but the testing that they've done all is non-specific; so, whatever people said.

Karen Pace: I'm quickly looking to see if they gave us any analysis of their exceptions under testing. Let me just look - exclusions, 2B3. And that's actually under 2B3 it says at the time of testing this

measure did not have exclusions. Or in their terminology exceptions which as they stated what they do is only look for exceptions or exclusions if the denominator condition was not met.

(Katherine): If we could just clarify, when this measure was originally endorsed it included a plan of care components. And that's why the measure at that time did not have any exceptions. We have updated the measure to be the pure intermediate outcome of the Kt/V level which is why now there are some examples with the exceptions.

We would be happy to - our clinical experts to define the residual renal function and provide that back to you. If that helps the discussion we would certainly appreciate that opportunity.

Peter Crooks: Yes. This was originally passed but not by the Renal Steering Committee. Right, Karen?

Karen Pace: No, this one was but it was that kind of crazy - it was a combination of the intermediate clinical outcome or if patients didn't meet that outcome then that they had a plan of care. So what they've done is they've dropped that plan of care and made it more a straight intermediate clinical outcome which I think is a good, you know, it's more consistent with the CMS measure. And it's less confusing of what it means to have a plan of care.

But I think what you're pointing out is there's some ambiguity about the potential exceptions.

Peter Crooks: Right. It could be game by the physician who's being measured very easily. So they could always have a 100%.

Alan Kliger: I guess my advice would be rather than to have us vote on this as it now stands. It would be wiser to ask the developer to take it back, to have a look at the comparable facility level measure that tech developed and to come back with a more precise measure for us to review.

Karen Pace: Okay. We can - we'll get back to the major developer and then let you all know.

Alan Kliger: Do others agree with that? Should we do that or? Yes?

Female: Yes.

Male: Yes.

Male: I do too.

Female: Yes.

Karen Pace: Okay, so we'll move on to the next one then which is, Lauren help me. 0253.

Lauren Richie: ((inaudible)) adequacy measurements of total fluid clearance at regular intervals.

Karen Pace: Okay. So this we're getting back to CMS facility measures. And these are comparable to the ones we talked at at the beginning about the hemodialysis. So the first one is about the frequency of measuring and then the next one will be about the method of measurement, so. This is percentage of all adults greater than or equal to 18 years old, peritoneal dialysis patients with total solute clearance for urea, and again in parens endogenous residual renal urea clearance plus dialytic measured at least once in a four month time period.

No exclusions, let's see, denominator, all adults on peritoneal dialysis process measures CMS and we can move on to talking about performance gap. So there didn't seem to be any issue with the preliminary evals. All rated high. Any question or concern?

Then we'll go on to evidence. And here we had once concern. Again, this will be the same issue we talked about previously. This would be indirect evidence. It's primarily about the relationship between the adequacy and mortality I believe.

And so I think some of the same issues you discussed for the corresponding HD measures apply here. But I'll stop there and see what the concerns about evidence might be.

Alan Kliger: Again, a tiny peripheral but as I wrote in my comment here. You know, Kt/V urea is currently the sort of standard measure of dialysis adequacy. And it's just no longer sort of the king of the hill. It's the only measure. Other measures of adequacy including volume measures time, KT alone without the etcetera, those are all being considered.

So when you look at the evidence. The evidence all has to do with this one particular measure. And it's reasonable again, we're passing on this particular measure. But I would just want to remind everybody that, you know, maybe too much of our focus has been on the urea kinetics.

And doesn't change our view of this, that's the only comment I'd make about the evidence.

Karen Pace: Okay. All right. So let me just ask if you have kind of the same approach or think it also applies as you were talking about the hemodialysis measures about if you think these should go forward. That during harmonization you may want to have a discussion about, you know, they should all be in some way in one measure with the adequacy.

Alan Kliger: Yes, indeed. I agree with you.

Karen Pace: Okay. So we can move on then and talk about reliability and validity. And the developer also I think this one is also in the - developer response with additional reliability information on this measure.

0253, the inter-unit reliability was .78. Okay. So I think - and I'll see if that was the issue about the reliability testing initially there was a question about the correlation over time periods. And then validity I believe they did the same analysis on this process measure to the standardized mortality ratio.

Alan Kliger: Yes. They did.

Karen Pace: Okay. And how did that come out?

Alan Kliger: Falling in mostly statistically non-significant performances almost surely for the same reason we talked about before.

Karen Pace: Okay. All right.

Any further discussions? Any questions, concerns? Okay.

Usability and feasibility on the preliminary values were rated mostly high from moderate. Any additional comments or questions? Okay.

And then overall suitability there was one concern probably related to the reliability and validity issues. So any other comments or questions that we should discuss with the group? Okay.

So then we can move on to 0254. This was - this is about the method of calculating Kt/V urea. It's the percentage of all adults greater than or equal to 18 years, peritoneal dialysis patients with a weekly Kt/V urea including endogenous and dialytic calculated in the standard way. And the numerator is - so this would be, I guess, the definition of a standard way.

It includes both the endogenous and renal urea's - endogenous renal urea clearance. And then residual function unless negligible which was defined as less 100 milliliters of urine in 24 hours. Assessed by measuring the renal component of Kt/V urea and estimating the patient's formulary filtration rates...

Alan Kliger: Very good.

Karen Pace: ...by calculating the mean of urea and creatinine in clearance. And then the third component is total body water or the volume, the V, estimated by either the Watson or Hume method using actual body weight. And BSA estimated by either the Dubois and Dubois method - that must be wrong. The Gehan and George - oh, Dubois and Dubois I guess two Dubois's.

Alan Kliger: That's right.

Karen Pace: Okay. The Gehan and George method or the Haycock method of using actual body weight during the four month study period.

So that's quite complex but I assume that - does that meet the standards for?

Alan Kliger: The answers, yes. These are the - sort of well defined - this is very nicely defined in fact. Precisely defined way in which is the way we do it.

Peter Crooks: Yes. Our RMA submitted should take a look at this metric in terms of specifying total (solid) of, you know, total Kt/V urea.

Because in the end we want all these metrics to harmonize into the, you know, using the same numerators basically.

Karen Pace: Okay. All right. So we can skip down to performance gap. Doesn't look like there was any issue in the preliminary reviews which we will get a clarification from the measure developer about all the data on the peritoneal dialysis because it looks like they cited those same numbers which is all patients as you...

Alan Kliger: Right. And also the discussion here is similar to the one we had for hemo which is that the measurement has two parts. One you have to do it and the second you have to do it in the standard way. And so there's 66% of PD, what we know, they're finding that 66% of PD patients are not - are not fulfilling this requirement.

It's likely that they're not having any measurements not that they're measuring it way different than is in this method.

Karen Pace: Okay.

Alan Kliger: So you know, similar - similar discussions to the other one.

Karen Pace: Right. Okay. Right. So we'll kind of pull in that discussion for both. Okay, any other comments about that.

Okay, evidence, it doesn't look like there were any concerns about the evidence. And we'll just point out as we've talked about before, just that this is indirect evidence. Or perhaps this is more direct evidence, again, it's about the method versus the actual value.

Alan Kliger: I mean that question about the economy. Their evidence doesn't deal with the fact that this is the best method. Their rationale had to do with the value of measuring (solid) clearance.

Karen Pace: Right. Okay. And there is evidence that this is the best method?

Alan Kliger: Well, I guess - this is silent on that question. I don't see any evidence for that.

Karen Pace: Right. Right. Okay. Okay. Any other comments or questions about the evidence.

Peter Crooks: Well, the - but the metric is actually, as you pointed out Alan, a combination of did you do it and then did you do this method. So I mean I guess there's evidence that it's good that you do it. At least indirectly but, you know, it does bother me that there's really evidence that their specific method - that they're specifying are better than other methods.

And, you know, however it's written that's sort of the way this will be taken. That things should be done by this method. But I kind - I think maybe under harmonization we can deal with this and, you know, if not this round. At least ultimately promote, you know, one metric that would cover all these bases.

Karen Pace: Okay. All right. So let's move on to reliability and validity. And again the developers response to the initial questions included a some additional reliability on the signal noise analysis and inter-unit reliabilities for 0254 was 0.64 and they did the same type of analysis for validity.

Alan Kliger: Similar to what we found in all the other PD?

Peter Crooks: Yes.

Karen Pace: Right. Some of the quintiles included one in the confidence interval which is probably again the small case volume issue. So any comments, questions about reliability and validity?

Alan Kliger: Just the same discussion we had before, you know.

Karen Pace: Okay.

Alan Kliger: Yes. Nothing new.

Karen Pace: Okay. Okay. All right. Usability and feasibility seem to - preliminary results rated high and moderate. Any additional comments or questions? Okay.

And then on preliminary assessment on suitable for endorsement there was one no and I think that was related to the reliability and validity issues which were reasonable questions. So we'll see how this shakes out. I think you've had pretty good discussion about that. Are there any other issues that should be brought up about feasibility for endorsement?

Okay. Lauren, what's next?

Lauren Richie: What do you want to talk about? Should we just table the harmonization?

Karen Pace: Right. We realized we missed sending you a couple of the measures specs related in competing measures table. But I think given the discussion today we need to get some clarifications before we're - before you're ready to think about harmonization. But are there any other comments about harmonization or competing measures that you want to highlight so we can note them when we do get to that.

Peter Crooks: Well, I guess you looked at two - at two issues. One is that if a metric is being submitted by AMA that is very similar but looking at a different reporting level that it match close as possible the way CMS has specified it.

And secondly that we want to look at reducing the hemodialysis metrics - facility level metrics - from three to one and the same for peritoneal dialysis if that's possible. I think that it is possible.

And that it would be more understandable and useful for public reporting to get it down to one - to one number that represents that the patients are being measured and meeting targets, so actually there's two - there's two different...

Karen Pace: Right. Okay and other committee members in agreement? Any other thing to add to that or other side of the coin?

Okay. All right. That's what I heard during your discussion so that helps.

Peter Crooks: So Karen, does this mean that every metric has now been looked at in that - by the committee in detail? We made it through all 34?

Karen Pace: Yes.

Lauren Richie: All except one which we're still trying to decide on that's the quality of life measure, so...

Peter Crooks: I was going to give ourselves a pat on the back but...

Karen Pace: No, and you guys, you all looked at that briefly at the meeting and discussed it. It's just whether they're going to be able to submit information for you to actually evaluate it. And I think there's a couple of other measures that, you know, we need to get some responses back to you all to see if you want to go any further.

But I would say essentially the answer is yes, you've looked at all 34 measures. And now we have all the follow up to do.

Peter Crooks: Right.

Alan Kliger: Can I just clarify what - where you're going to put 0321? How are we going to handle 0321?

Karen Pace: Okay. Let me just get my notes here.

Alan Kliger: That's the one that, you know...

Karen Pace: Right. So my understanding is that we want to ask the developer for clarification about the endogenous kidney function in the numerator and then also about having precise specifications about it in the exceptions, the residual kidney function as well as the other medical reasons.

Alan Kliger: Okay.

Karen Pace: And the suggestion is for them to take a look at the CMS measures about potential ways to define those.

Alan Kliger: Particularly 254.

Karen Pace: Okay.

Alan Kliger: So we will defer further consideration of that measure until those - that resubmission - is that what I understand?

Karen Pace: Yes.

Alan Kliger: Okay. Thank you.

Karen Pace: Is that, unless I - is that?

Alan Kliger: Yes. That's what I was...

Peter Crooks: That's our preference.

Alan Kliger: ...listening for but, yes.

Karen Pace: Okay. All right.

Lauren Richie: And I'm only going to ((inaudible)).

Karen Pace: All of them ((inaudible)) so after this - go ahead.

Lauren Richie: Yes. So after the call, just as we've done the last two workgroups, you'll be getting a link from me shortly, another survey monkey link to submit your votes, again, based on just today's discussion on the measures. All except 321 that we just talked about. And we'll have to send this up to plan survey after we get more information from the developer on that.

Alan Kliger: Okay, great.

Karen Pace: I think, Tom, for our member and public comments period. So operator if you can please open all of the lines.

Operator: Absolutely, those lines are now open.

Karen Pace: Okay. If there are any public comments or comments from the ((inaudible)) numbers on the line? Any comments from the developers?

Okay and from the Steering Committee members any suggestions to us at staff on moving forward with these and next steps or any other suggestions are welcome.

And then, also, as a quick reminder we still have two more calls left. And hopefully that's it. We did identify an opportunity to discuss some (departmentalization) issues with the anemia and cardiovascular measures. So that call will be October 4. I'm blanking on the time I believe it's...

Kristine Schonder: 3:30 to 5:30.

Karen Pace: Thank you Kristine. And then our final call on October 13. I believe that is 2 to 4 pm Eastern. So two more calls and I'll send out reminders for those as well.

So I guess we could have a brief discussion or see if you have any - what our thinking is that we will get all of these workgroups final discussions and ratings need to go back to the full committee to make a final judgment on these measures. We were thinking that one way to do that as efficiently as possible is to get that information back to you. And then have the Steering Committee vote on line. And then we would discuss the results of that voting on that last call.

Because, again, that's two hours and, you know, we would have a hard time getting through all these measures - these final measures. There was one suggestion made that when we do that with full Steering Committee to do it by topic area. If you have some other suggestions or totally new suggestions of how, you know, we should consider proceeding with the full committee once these workgroup calls are done we'd be glad to hear those.

Alan Kliger: So I'm sorry, just help clarify for me. Your idea would be to send out the preliminary workgroup ratings and comments to everyone and ask everyone to vote and then discuss it afterwards?

Karen Pace: Yes. By topic area.

Peter Crooks: Yes. I would think that also should include your summary of today's call. You know, the summaries...

Karen Pace: Yes. Yes.

Peter Crooks: ...you've written up of the - at least I saw the one from the first call. It was distinct, it really touched on the key issues that we spent time on and summarized them very nicely.

Karen Pace: Right. Yes. Yes. It would definitely have to be with a summary of the workgroup discussion and their voting. Now if that doesn't - the other option would be to have the call and have the discussion followed by voting. We'd have to figure out how to really make the discussion kind of laser in on any issues. And I'm not - so, you know...

Alan Kliger: Well, here would my only thought or suggestion. I like your idea, it's a good idea. I wonder if there's a subset of measures for which after all of these processes are done there is substantial disagreement in the workgroup and if those might somehow be called out and people would be treated to an in person discussion. Focused but, you know, clear before they actually voted.

Karen Pace: Right. Okay.

Alan Kliger: So I mean of all the measures I don't know, how many did we look at? How many we going to be voting on? It's probably going to be something like 20 of them?

Karen Pace: I think about 17, yes.

Alan Kliger: Yes, 17, if, you know, 13 of 17, you know, it's either unanimous or one person didn't agree in general to the assessment. To take the remaining two or three and asking people not to vote those ..

Karen Pace: Sure until we have...

Alan Kliger: ...over the line until we have some discussion about those.

Karen Pace: Yes. Well, that makes sense.

Peter Crooks: Yes. I think that makes sense. I guess we could meet, Kristine and I and Karen and Lauren, sorry, Lauren and Karen, right. And maybe step through the 17, 18 measures and see if any would - we think really qualifies as still quite controversial and needing discussion. And how many of those there are.

Karen Pace: Yes. Yes.

Alan Kliger: There may be none which would be easy. Or there may 12 which would be impossible.

Karen Pace: Right. Right.

Alan Kliger: But I would, yes, I think that your idea of looking at it that way might be smart.

Peter Crooks: I do like the idea of for the majority at least of the metrics to have people vote on line. I think you'll get a more thoughtful process. People have to pull out the materials and read it in their voting. Where if you kind of doing it in a two hour meeting a member could come kind of unprepared and just kind of go with the flow and not give their best effort. So I do think it's

worthwhile to have people thinking about it off line. Reviewing what the - these sub-committees have gone through.

Karen Pace: Okay.

Kristine Schonder: I agree and I think I would kind of outline that in the instructions for the Steering Committee as well especially for those members who aren't part of any of the workgroups.

Karen Pace: Sure. Okay. All right. Well, you know, if you think of anything else feel free to send an email to Lauren and myself and we'll continue on. We'll be getting out the summaries from these two calls as quickly as possible. But Lauren will send you the voting tool while it's fresh in your mind. And we'll follow up with the summaries and we'll proceed.

We appreciate all the time and effort of going through these measures and the time and effort of everyone on the call. Lauren, anything else?

Lauren Richie: That's it. Thanks again, everyone.

Peter Crooks: Okay, thank you.

Alan Kliger: Okay.

Male: Thank you all.

Female: Bye.

END