NATIONAL QUALITY FORUM

Moderator: Alexandra Ogungbemi
April 21, 2015
1:00 p.m. ET

OPERATOR: This is Conference #: 25691673.

Poonam Bal: Welcome everyone to our second workgroup call. If you were not able to join us on the first one, we'll walk you through the structure of how to discuss each measure. I'll kind of lead you through the first one and we'll move forward with that.

So, I did want to start with being who's on the line and, Alexandra, if you could just do a real quick roll call.

Alexandra Ogungbemi: Good afternoon. Is Ishir Bhan on the line?

Ishir Bhan: Yes, I'm here.

Alexandra Ogungbemi: Thank you. Peter Crooks?

Peter Crooks: Present.

Alexandra Ogungbemi: Michael Fischer?

Michael Fischer: Present.

Alexandra Ogungbemi: Mahesh Krishnan?

Mahesh Krishnan: Present.

Alexandra Ogungbemi: Lisa Latts? And Joshua Zaritsky informed us that he would not be able to join the call today. So, thank you.

Poonam Bal: Yes. So, unfortunately both Lisa and Joshua will not be able to attend. Since they also team together for two of the measures. I'll go ahead and ask in advance if anyone wants to take the lead on those measures, if you feel (inaudible) to start the discussion, but we'll get to that when we get to these measures.

So, as everyone knows, we're trying to arrange these workgroup calls so you're familiar with the process that we'll take in the in-person meeting. So, we'll start out by asking the lead discussant to do a quick overview – a quick – brief introduction to the measure. Then (inaudible) again. And then once the committee had their opportunity to ask any questions or make any statements, we'll move on to the first category which is importance, that is broken up into multiple categories. So, we'll drop down to evidence first and we'll discuss evidence.

The lead discussant will start that discussion. We'll open up to the committee and allow them to do questions and continue to move through that way. And that's how we'll do it in in-person too.

I will be acting as how the chair will be acting in the in-person meeting. So, I'll be the one kind of directing us through that and letting you know when to move on to the next section or maybe something you're discussing should be delayed until we get to that specific (in**audible**).

And so, with that said, I want to see if there's any questions about how this would be structured before we start.

Peter Crooks: This is Peter Crooks. I have a suggestion. I don't know if this makes sense to you and the others on the group and would it make sense to order these such that we do the two hemodialysis adequacy measures back to back followed by the two P.D. measures and then the other two laboratory measures. Is that – it seems to me as long as we have our heads into the hemodialysis adequacy, we can move from one to the other.

Poonam Bal: OK …

Male: Yes.

Poonam Bal: Does that sound …

Male: I like that.

Peter Crooks: So, I would suggest your order would be like 0249 followed by 0323, and then we could go to 0318 and followed by 0321, they are in order, and then do the phosphorous and then the calcium measures. So, that would be 0255 and 1454.

Poonam Bal: OK, perfect. Thank you for that suggestion and we'll go ahead with that method.

So, let's see here. So, 0249 is first. So, Peter and Ishir, if you want to decide who would like to go first and start the discussion starting with the brief introduction of the measure.

Ishir Bhan: I can sort of introduce this one. So, this is the delivered – 0249, delivered dose of hemodialysis above minimum from CMS. And the brief description here is the percentage of all patient months for all – for those patients who are basically greater and equal to 18, whose delivered – average delivered dose of hemodialysis calculated from the last detriment of the month was between – and the lower threshold here of S.P. Kt/V of greater than or equal to 1.2 and the upper threshold of 5.0.

And the rationale here was that the public studies indicate that there's an association between low S.P. Kt/V and increase in mortality low here being less than 1.2. And the 2006 KDOQI guideline suggest that a minimum adequate dose of three times weekly dialysis should be above that threshold. And furthermore, they add here pediatric population should also receive at least a S.P. Kt/V of 1.2 which is the minimum adult requirement.

So, just to move on, the numerator here is the number of patient months and the denominator, whose delivered dose hemodialysis actuate from the last

measurement of the month was between 1.2 and 5. And the denominator, a patient must be 18 – greater than or equal to 18, must have ESRD for greater than 90 days, must be on three times weekly dialysis during the month and must be assigned to that facility for the entire month.

Exclusions here are pediatric patients less than 18, those who have less than three times weekly dialysis, patients without ESRD for less than 91 days and patients in the facility for less than a month. So, they're basically all the same but are not included in the inclusion.

So, I'll move down to the evidence here. And what they – it's an intermediate kind of outcome measures and the evidence, to demonstrate that an average delivered dose of hemodialysis between 1.2 and 5 is linked to the reduction in (undesirable) health outcome. And we just go down a bit here. Basically they say the clinical practice guideline presented from KDOQI 2006 or through – already stated earlier, were just that minimum delivered dose of H.D. three times per week, it should be at least 1.2 per session. And that's graded A strong evidence.

The developer here provides a summary with 5 studies study that measure dialysis dose doing a single-pool of Kt/V and one was a randomized clinical trial. That's the HEMO study with 18 to 46 patients. One was a prospective study with 140 patients and the remaining retrospect study with sample sizes of 78-71 and 11-51. And two of the studies found a significant improvement in mortality with increasing S.P. Kt/V. The remaining study compared higher doses of S.P. Kt/V in the standard dose which is the threshold, the 1.2 and found that the higher doses were not beneficial in improving mortality compared to the standard dose.

OK. And so just for the folks wanting the call, how do you want to proceed here? Should I just keep going down or …

Peter Crooks: So, I think you should pause here and then, Poonam – then everybody has a chance to discuss evidence before we move into the performance gap.

Ishir Bhan: Got it.

Poonam Bal: Yes, that's exactly right. And before we do start the discussion, I just wanted to provide some reminders. For those that are on the webinar, the pink section you're seeing right now is a summary of the evaluations done by the committee before this meeting. So, one familiar way that you can find the results of your – the analysis of your colleagues, this is in the measure worksheet and it'll be in a pink box. That's one reminder.

And then also, as we start the discussion, I want to remind everyone that we do have developers on the line for the purpose of answering any questions you may have. And we can call up on them when necessary. But, otherwise, the discussion is mainly between the committee.

With that said, I'll go ahead and open up for discussion on evidence.

Peter Crooks: This is Peter Crooks. I was the second primary reviewer on this one. And in terms of evidence, I – you know, it is old. I mean, it's – they're basically KDOQI 2006, all the references are pre-2006. However, I think it, you know, probably still holds up. They were kind of sloppy, I have to say, in filling out the evidence review form and now having been a developer, I understand why it's challenging. But, you know, I think it does call attention when you just leave certain parts blank or you answer inappropriately, in one case, answering a different question than the one that's asked.

Also, the – in last section of the evidence review, they asked for any – are there any recent publications, and please cite them and give us some infringed one. They list four papers at some of which I'm not even sure they're really research studies and they did not give a summary for each one. So, again, it's not being responsive to the questions. If this – If passing depended on filling out the form properly, this wouldn't pass. But I have to say I believe the evidence from the KDOQI work in 2006 was based I believe on a systematic review of 10 or 12 papers and I think it holds up.

So, that's my comments. Thank you.

Ishir Bhan: Yes. Pretty much thought the same way that, you know, the – for what its worth, the evidence seems pretty reasonable for looking at the lower threshold, 1.2. My sense is that, there wasn't much guiding the upper

threshold but I thought there was many patients who are above that threshold regardless.

Peter Crooks: Yes, why an upper limit of 5.0 except to maybe find data entry errors.

Ishir Bhan: Yes, maybe that's the only thing I thought.

Peter Crooks: But that's somewhat arbitrary though, correct?

(Crosstalk)

Peter Crooks: … evidence presented for that number.

Ishir Bhan: Yes, I think just arbitrary. And I think it's – I didn't see any logic provided for why you do an upper threshold. But on the other hand, I can't imagine those many people who are in that range unless it's an error.

Michael Fischer: Yes. This is Michael Fischer. Yes. My only thought is why have an upper limit at all? If we're going to have an arbitrary one that's not based on any evidence and it maybe begs the question, why have one at all.

Mahesh Krishnan: Yes. I mean, what – this is Mahesh. When we hear this as a distribution, it's a standard distribution, right, so (bell) distribution. So, I'm with you. I don't think the upper limit makes a lot of sense.

Poonam Bal: At this time, I want to interject and see maybe the developer would like to respond. If anyone on from the University of Michigan or CMS?

Joseph Messana: Yes, this is Joe Messana from University of Michigan (CAC). So, I appreciate the comments of the reviewers. There are a couple of reasons why the – why a plausible threshold was included.

In the original submission, a plausible threshold of 2.5 was included from the depths of history related to the kidney data dictionary and other definitions. And some public comments from DFC preview I believe from last year pointed out that some patients who were on nocturnal in-center longer dialysis were achieving Kt/Vs in the 2.5 range, a little bit above and a little bit below.

And so, that was one reason why we arbitrary – arbitrarily chose and a plausible threshold was increased to 5. It turns out there were very small number. I don't know the exact number, but I believe under 100 or under 150 values that are above that threshold. However, there are two data entries that are numerical values from the claims, Kt/V reporting 8.88 and 9.99 that can be submitted to identify – one of them is missing data and the other one I don't recall right now, but there are two placeholder numbers.

And so, we would have to exclude – at a minimum, we have to exclude those two numerical values. And then the question is if people are uncomfortable with Kt/V calculations for single-pool that are above 5 which reflect extraordinarily high urea reduction ratios basically.

And so that's the only justification. Thank you.

Mahesh Krishnan: Joe, I didn't quite catch that. Are you saying that the upper limit of 5 proxy was due to pro-hemo as opposed to just in-center hemo? And then I'm assuming the 8.88 and 9.99, those are part of a code for micro specifications not part of the measure, right?

Joe?

Male: It seemed that we have lost him.

Poonam Bal: Is there anyone from the University of Michigan or CMS that will be able to provide the answer?

Yes, I think we lost the developer for now.

(Tammy Sharon): Sorry. This is (Tammy Sharon) at the University of Michigan. I think the question about either it was home hemodialysis, there had been comments to DFC saying that people are nocturnal dialysis would have, might go higher values of Kt/V and that those are legitimate values which is why we got at that upper limit needed to be higher than it had been before.

And I'm not sure if that answers the entire question.

Joseph Messana: Yes. And I apologize. This is Joe Messana. I got cut off. So, I'm back and I missed the last minute or minute and a half of discussion after my statement.

Mahesh Krishnan: There are two questions, Joe. One was – So, Mahesh. One was the upper limit for nocturnal hemodialysis, I guess you're counting that as an in-center modality for this. And then two was with the 8.88 and the 9.99, those seem to be more micro specifications that you would program out, right, as an exclusion as opposed to part of the measure, rather than being a rationale for the Kt/V of 5.

Joseph Messana: Yes. But the original intent was to prevent entry of implausible values I think, as (Tammy) was mentioning. And it turns out that the old threshold which was developed before people were doing long nocturnal or long dialysis rarely if ever would be achieved with a single-pool Kt/V which is what we're talking about. That's all.

Mahesh Krishnan: OK.

Joseph Messana: OK. Thanks.

Poonam Bal: And so, I just wanted to – we have received an answer and then we'll be able to discuss this further in in-person meetings, but I wanted to see if there was any other points that people want to bring about evidence before we move forward.

Mahesh Krishnan: Just a question when we think about this for the group, which I think Joe just raised which made a lot of sense to me. The evidence base here is in in-center hemodialysis, right, three times a week hemodialysis. Are we considering alternative in-center modalities whether it's – I guess it's mostly nocturnal would be the other in-center modality – as part of the same evidence base or does that require different evidence base? I'm just trying to think about that on my head.

Peter Crooks: Well, the measure only addresses patients that get – this is Peter – three times a week.

Mahesh Krishnan: Yes.

Peter Crooks: So, I don't think they need to address other issues except for that.

Mahesh Krishnan: Yes. Other modalities, right?

Peter Crooks: Right. Or other schedules.

Mahesh Krishnan: Yes. OK.

Peter Crooks: OK. Should we go to gap?

Ishir Bhan: Yes, why don't we do that? So, the performance gap here …

Poonam Bal: Sorry, to interrupt.

Ishir Bhan: Yes, yes.

Poonam Bal: Just one more question, I'm sorry. I'm one of the NQF staff here. And just so everybody has practice going through the algorithm that sort of thing, we did put it up on screen. And I wanted to go back to one question. You had noted that the guidelines were a little bit old, so I think I have two questions for the committee to make sure that everybody is on the same page.

Number one, is there anything new or in terms of guidelines that you're aware of that's newer than the 2006? And if not, are you guys sufficiently familiar with the references that the developer cited, the newer stuff, or do we need to ask the developer to very quickly summarize that and then you (inaudible)?

Peter Crooks: The newer stuff that is mentioned in the evidence is really I think almost editorial type stuff not really evidence. I'm not aware of other systematic reviews or major evidence on this. But I'm not an academia. Do others have – aware of other good evidence?

Ishir Bhan: I'm not aware of anything major that's come along since then.

Mahesh Krishnan: Yes, agreed.

Poonam Bal: OK, great. So, we just want to make sure that everybody is comfortable with the evidence algorithm. It's on your screen there. So, if you guys want to talk

this really briefly about how you would work your way through the algorithm and where you might land in terms of a potential rating. Again, this is just to give you practice to go through that.

If everyone feels comfortable using your own, we can move forward as well. But we do want to make this as effective as possible. So, if people are not familiar with the algorithm and would like to go through it or just want a quick recap from our measure Q&A (inaudible).

Mahesh Krishnan: This looks pretty straightforward to me.

Ishir Bhan: Yes.

Poonam: OK. So, just I think a couple of things to point out, usually guidelines often are evidence-based and can be considered systematic review. And it does sound like the developer, if I understand correctly, did give you a summary of the QTC. So, that would put you probably about four for and then on to boxes 5 A, B and C and that would help you land on your rating.

OK. I think you can go ahead right now. Ishir …

Ishir Bhan: OK. So, moving along, and let me just load up the next section there.

So, performance gap, performance, you know, it's based on the 2013 CROWNWeb and Medicare claims data. This advanced 2013. We looked at 55 – 5,557 facilities with at least 11 eligible patients. And this was striking, the mean performance score was 93.5 percent with a standard deviation of 7 percent. They do provide some disparity data as well. And the developers stated that in the absence of biological effects explaining the difference of those risk adjustments, risk factors would potentially mask disparity in care.

And so, the questions for us here were with the mean performance of 93.55 percent is our gap in care that warrants a national performance measure. And I think this is the most striking for me if that was starting off with a pretty high score. And I'll let others chime in.

Peter Crooks: Yes. And this is Peter Crooks. I agree with that concern the – and you think about 7 percent or what was the gap? 93.5 percent versus 100 percent, that's a 6.5 percent gap. And we know in the real world of giving dialysis, it's pretty hard for every patient always to hit the mark. They may have catheter problems, they may have flow issues. There's always going to be a small percentage of patients that don't hit the mark. So, taking that into consideration, it's conceivable that the dialysis world has pushed this up to about – this metric up to its max. And if that's the case, should it be continued as a performance measure or not is a valid question.

Michael Fischer: So, this is Michael Fischer. I mean, I agree with all of the comment. I think in one of the documents, NQF addresses this particular issue about is the performance "topped out." And I guess one also needs to think about if something is no longer endorsed, what are the risks of losing endorsement. So, I guess that's something the committee, we just need to kind of consider.

And I guess the other option in such situations is reserve status.

Peter Crooks: Yes. Could the staff talk about that? Is that an option for us to recommend?

Poonam Bal: Yes. Absolutely. So, if it turns out that you've judged the measure to meet all the criteria except for gap and you think that it's topped out, then basically we would consider it for reserve status.

So …

Peter Crooks: What does reserve status really mean?

Michael Fischer: Yes.

Poonam Bal: So, a second. I'm looking at my policy here. Measures assigned inactive endorsement status or reserve status won't be reviewed in the usual maintenance review cycle. However, the standing committees would periodically review measures for any change in evidence or deterioration and performance or unintended consequences or any other concerns. So – And the committee could also remove a measure from inactive or reserve status if the measure no longer meets the endorsement criteria.

So, basically what it means is it's still a good measure, it's just topped out. And it just means that it's there, it's not making any statement other – you know, it's not saying that something is wrong with the measure, if that makes any sense. But it's also saying that there's – it probably has done the most that it could in terms of having an opportunity for improvement.

Peter Crooks: So, it remains an endorsed measure …

Poonam Bal: Yes.

Peter Crooks: … and the – whoever is looking at it, the industry could say, "This is an NQF endorsed measure," even though we're doing, well, we want to keep our eye on the ball. And then you said the standing committee or the NQF will review it again in three years or at some point in time to keep assessing its status.

Poonam Bal: Right. And let me read you the actual piece here while – if a measure does go into this reserve status, it would communicate to potential users that the measure is no longer addressed high leverage areas for accountability purposes. So, they are retained in our portfolio, they are endorsed, the committee would look at it periodically just to make sure that it's still meeting the criteria for endorsement. But you're sending the message that there's really not as much need for that measure anymore.

Peter Crooks: Great. Thank you.

So, at our in-person meeting, we can discuss this further as related to this particular measure.

Poonam Bal: Right. And it's a little tricky because you talk about the topped out and then you also have to make sure that the other criteria are still met. So, again, reserve status only makes sense if everything else is met except for gap.

Peter Crooks: Right, right. Which may be the case.

Poonam Bal: Right.

Peter Crooks: OK.

Poonam Bal: OK. So, how about we move forward to the second criteria, scientific acceptability starting with specifications?

Ishir Bhan: OK. So, specifications. Reliability specifications. It's two – This is 2A1. And this measure assesses the number of patient months and the denominator would be delivered dose of hemodialysis is between 1.2 and 5. The denominator includes patient 18 years or older who've had ESRD for greater than 90 days, dialysis three times a week and assigned to a specific facility for the entire month. Data elements are defined and derived from treatment history file that's available for all patient and dialysis. And the developer indicates that this measure is collected from CROWNWeb and if data are missing, Medicare claims are used. There's no mention on the need for ICD-10 or CPT codes with measure and be an intermediate outcome measure with no risk adjustment.

So, first question is, are all data elements fully defined, are all appropriate codes included, is the logic or calculation algorithm clear and is this – likely this measure can be – can this be implemented. My thought is this is all pretty straightforward and pretty logical. This data easily pretty readily available, but I'll let everyone else chime in.

Poonam Bal: OK. Well …

Mahesh Krishnan: Yes, I …

Poonam Bal: Sorry, go ahead.

Mahesh Krishnan: It's Mahesh. I have two questions or two comments related to the data on the specification. One, my experience has been that for many of these, the statements that the patient is at the facility for a month is also a bit crude. It would be better to define as more clearer, I think in other measures specified as having about seven treatments per month and that comment will echo through.

Two, just in terms of where the data comes from. I'm the executive sponsor for CROWNWeb for DaVita and I can tell you that today we have anywhere

from a 5 to 15 percent data loss in terms of the data transmission. So, the claims actually today have a higher fidelity than CROWNWeb does today. Everyone is working on solving the problem. But because of the matching algorithms that are used in data transmission, even though there's valid data, it just doesn't get into the system.

And so, those are just my two comments.

(Off-Mike)

Poonam Bal: … comments?

OK. And we can move forward to reliability testing.

Ishir Bhan: OK. So, on to the next section, reliability testing. Testing was performed at the measure level using data from 2013 CROWNWeb and Medicare claim. 5,557 facilities that has at least 11 eligible patients recruited in the analysis and it's amounted to 368,392 patients. Developer calculated the IUR for the overall 12 months to assess reliability of the measure, IUR measures the proportion. The measure variability that is attributable to the between facility variance.

And the developer indicates that the one way (innova) recommended by the NQF should be interpreted with caution if method is appropriate for measures that are approximately normally distributed across facilities, but this measure is not normally distributed. The IUR is 0.4 – sorry, 0.942 which can be interpreted as high and suggest that 94 percent of the variation of this measure is attributed to be between facility variation.

And our – questions for us are specific questions of the method and results of reliability testing, is the sample adequate to generalize for widespread implementation and do the results demonstrate sufficient reliability so differences in performance could be identified.

Peter Crooks: So, that's the standard approach that we're seeing in a lot of these to measure the inter unit variability and to show that it's higher than – the intra is higher than the inter unit variability. It seems to meet criteria to me.

Poonam Bal: Were there any additional concerns, committee?

OK, then we can move forward to validity. Thank you.

Ishir Bhan: So, with validity specifications here, this measure is based on studies indicating an association between low S.P. Kt/V and increased mortality. In addition, the KDOQI hemodialysis adequacy guidelines specified that minimal adequate dose of hemodialysis given three times per week to patients with a K.R. or less than two notes per minute over 1.73 meters squared should be an S.P. Kt/V of 1.2 per dialysis. This measure specified to assess performance of dialysis facilities in achieving optimal care based on evidence and guidelines.

Are these specifications consistent with the evidence? It seemed to be for me.

Poonam Bal: OK. Peter, did you have any additional comments before we open up to the committee?

Peter Crooks: This is Peter. Yes, I'd say they have met the task of linking the outcomes to mortality and also have claim faith validity based on a clinical TEP group. I'd give it a thumbs up.

Poonam Bal: OK. Was there anyone on the committee that wanted to make any comments?

So, then, we can move forward and start our next discussion to feasibility.

Peter Crooks: Going further down into this section, they had the opportunity to compare the Medicare claims data to the CROWNWeb data to show that they're equivalent data sources and they didn't do that. I think that would strengthen that if they're going to be kind of going back and forth, you know, which data source are they using, it's good to show that you're getting equivalently – equivalent data off of that.

And then another metric by CMS, they did do that analysis. I forget which metric it was, but they didn't do it in this case. So, we just have to kind of take on (face) of it's fine to do it from claims data or CROWNWeb.

Ishir Bhan: Yes.

Peter Crooks: You know, maybe the developer could address that at the in-person presentation.

Ishir Bhan: OK. Could we move on to feasibility? The data source for this measure is CROWNWeb and if the patient data is missing in CROWNWeb, then Medicare claims are used. They're just used for backup (data). 5,557 facilities had at least 11 eligible patients during the 2013 period and were included in the analysis. Data is collected or generated and used by health care personnel during provision of care. Excuse me.

The questions for us, are the required data elements routinely generated and used during care delivery, are there prior data elements available in electronic form, (ET) from EHR and other electronic sources, and is the data collection strategy ready to put out for external use, and is an eMeasure does – and eMeasure (inaudible) scorecard demonstrate accessible feasibility and multiple EHR systems in sites.

So, just for me, the data collected during the routine delivery of care and did appear to be readily available. So, I don't see any major barriers here.

Mahesh Krishnan: And Peter, just Mahesh again. I just can't stress this enough, guys. The – Like I – The data is available but the amount of data that gets transmitted through CROWNWeb is not 100 percent versus in claims, it's nearly 100 percent. So, back to Peter's question, I think we just really got to look at the data source and whether or not CROWNWeb is actually ready for a primetime in terms of utilizing this data, especially the metric that just topped out, if you have data transmission issues, that can reflect – that can create an artifact in terms of unit performance that's differential from what actually happens in the EMR which you don't want.

Poonam Bal: So, with that said, I do want to give the developer an opportunity to respond to that. If anyone from the University of Michigan or CMS would like to respond.

Joseph Messana: Well, so, this is Joe Messana. I think that it's a clear definition of the determination of patient denominators is in order here. Our treatment history file is for people like Mahesh and the few – some others who may be familiar with how the USRDS calculations, placement of the patients in facilities at first is done in a similar fashion.

So, we start with the – essentially the old (SIMS) extract from the CROWNWeb data and then we supplement that with claim – Medicare claim status who place patients in facilities that have been – have a discharge event but don't – but that hasn't been completed in CROWNWeb. So, we pick up 4 or 5 percent of additional patient placements in facility. And so that's the treatment history description here.

It should be noted that that is a different methodology than the historical claims-based metric that people are familiar with from some of the public reporting types like DFC. And the intent here to do that is to reduce the issue, the problem with transient patients and to use CROWNWeb but is that – maybe incomplete, we supplement it with Medicare claims files.

So, at least for patients who have Medicare primary or secondary claims, the CROWNWeb data are being supplemented. That's what the denominator is that we're proposing. And the numerator similarly we're utilizing the single-pool Kt/V data from CROWNWeb and if it's not present, if it's missing but we have a patient in the denominator with Medicare claims, we're using the value from Medicare claims to try to make it as complete as possible.

And if you'll notice, if you look at the rate in the 93, 94 percent moving the threshold for adequate and you compare it to national data from DFC from a year ago which was in the 88 percent range, most of that difference is because of this new methodology that's been proposed using the treatment history file. So, it eliminates many of the missings which I think many in the community I believe are due to transient patients and may not fairly represent a facility oversight.

Mahesh Krishnan: Sorry, Joe. Just to be clear, are you saying this metric will differ from the prior metric because you're going to use (SIMS) data and CROWNWeb event

data to allocate where a patient is or is not as opposed to a minimum treatment count which is what was being used before like seven treatments in a month?

Joseph Messana: Well, minimum count was – the minimum treatment count has not been used other than the two – as you probably know, (Cliff) used the two-touch rule. So, if you had more – two or more treatments in a month, a patient was assigned to your facility, for the public reporting measures, it was a one-touch rule, one or more treatment. This metric includes a patient in a facility if they have a CROWNWeb admission event or if they have extensive claims history for the month that meet a conservative standard for that.

And we can provide a bit more background information for anybody that would like, but for the claims – excuse me, for the CROWNWeb-based treatment history approach, that fundamentally changes things and I think increases the patient attribution accuracy compared to the one-touch or two-touch historical approach in the previously endorsed metric.

And I will – I could go on for a while about that. But – So, fundamentally, though, if someone chose to calculate this solely from claims, their advantages and disadvantages that only includes Medicare patients. One would have to sort out the number of touches that would be appropriate to consider attribution to facility. As Mahesh has pointed out, the number seven that's been raised.

Mahesh Krishnan: Yes, it would be helpful to get more data on that, Joe. I did appreciate that level of complexity but it would be helpful to understand that because even on the admission of discharge records for CROWNWeb and also errors where they're miscounted and then have to be fixed, and that's not a thing that CMS is aware of and that we're dealing with.

So, it would be – I think it's important just to understand the nuances and reflect that against the current reality of what the known problems are with CROWNWeb today.

Poonam Bal: All right. Well, thank you so much for that discussion. The committee should definitely keep that in mind going through the other measures, very similar data sources.

If there were any other comments on feasibility, please let us know now, otherwise we'll move on to usability and use.

Ishir Bhan: So, for usability and use, currently reported in the dialysis facility compare public program in which all Medicare certified facilities eligible and have at least 11 patients are accountable. According to the latest DFC report, that was 5,625 facilities. It's currently reported in the ESRD QIP payment program in which all Medicare certified facilities are eligible and have at least 11 patients are accountable according to the last DFC report that 2,640 facilities. And it's planned to be used for quality improvement with benchmarking.

So, the questions for us, is the measure publicly reported per maintenance measure, is the measure used in at least one accountability application, how come the performance results be used to further the goal of high quality interest in health care and do the benefits of the measure outweigh any potential unintended consequences.

So, it does look to be publicly reported.

(Off-Mike)

Peter Crooks: Now, this passes the usability criteria in my view. This is Peter.

Ishir Bhan: Yes. Agreed.

Poonam Bal: Any comments from the committee?

OK. Well, that concludes our first measure. We've done it in a record time. So, thank you for Peter and Ishir for your great guidance.

Just some – Those are the general process of how we'll be discussing and eventually building on the measures in the in-person meeting. When we would pause after each period, but when you vote, obviously (inaudible) done. The role that I was taking on would be generally the role of the chair. They'll be guiding us through the discussion. The developer would be providing the initial summary of the measure and then the lead discussants would be the ones starting with evidence and going through that process.

So, those are generally the same process. Some areas on the worksheet that I did want to point out was at the end of this, you'll also see that there's a related and competing section. We do not discuss that right now but we have asked the developers to provide responses and this is one response. And we'll discuss that during the in-person meeting.

Also, underneath that, you'll see that there's a pre-meeting public and member commenting. This measure does not have any comments. But if any comments did come in for a measure, they would be in this section. It's in that purple color. So, we do ask that you look through those as well before the in-person meeting.

And I will move on to the next measure which is 0323. And the lead discussants on that, we have Mahesh and Michael Fischer. Before we start, you know, keeping time in consideration, we are – that's the general process and that's how we'll be running the in-person meeting, but to make sure that we give each measure enough time, we're going to be doing it slightly different where we'll ask the lead discussants to give a brief description of the measure and then instead of category by category, to discuss just the major issue that they had. So, anything they want to pull out to discuss, we'll focus on that instead of going through each criteria.

Again, that's not how we'll do it in the in-person meeting but we do want to make sure that each measures get the time that they need.

So, I'll start with Mahesh and Michael Fischer, if you guys want to decide who wants to go first.

Mahesh Krishnan: Yes, Michael, it's up to you.

Michael Fischer: Please, feel free, Mahesh. Thank you.

Mahesh Krishnan: So maybe just a summary and I'll turn – I'll take your guidance through most times. So, this is the – I look at this, as Peter mentioned earlier, the companion measure to the prior measure that was being discussed, this being a physician level measure as opposed to a clinic level measure. And the brief

description of the measure is the percentage of calendar months for the 12-month period during which patients 18 years or older with ESRD receiving dialysis – hemodialysis three times a week for greater than 90 days, have a Kt/V or greater than or equal to 1.2.

And the rationale was very similar to what we just talked about before, so I won't spend a lot of time on that. There were no exclusion – no denominator exclusions or exceptions for this measure. When I went through this measure – so you want me just to go through the highlights of this measure then, right?

Poonam Bal: Yes please.

Mahesh Krishnan: So, when I went through this measure, it was actually very similar to what we just talked about in terms of what made sense and what didn't make sense to me. The evidence to support the measure, I thought there were sufficient literature to do so. And the performance gap again was based on older data. It was 2008 data, and in 2008, it was reported that 41.36 percent of patients did not receive optimal care. That's pretty old data. But if we look at the data we just talked about in the 0249 measure, that gap was significantly lower.

So, I assume that that smaller gap which exists in the clinic level data should be similar in nature to the performance gap for the physician level measure perhaps because it's a yearly average, we should expect the payout to be higher percentage. But I think that the gap should tell all that from the newer data.

And given that there is a physician – this is a physician level measure, it's possible that gaps still exist. The data submitted for disparities indicated there was a gap in access to hemodialysis by race, but that was old and didn't really answer the question around adequacy by race. But when we look again at that parallel measure, there was statistically significant issue by race, so that was raised and I guess that could be statistically adjusted for.

That was a high priority area for me because they (also had) adequacy, in my opinion, as we talked about the early measure, is a fundamental criteria for quality and dialysis which has affected both by the clinic as well as the physicians writing the prescription and there was no composite measure.

And that was the end of section one. I don't know if, Mike, do you want to comment on that or I can go through the other section just all at once?

Michael Fischer: I agree with all of your comments. Just very quickly, one I had concerns, the data is old, in 2008. And I reviewed this application on its own ignoring, you know, what we had discussed in the prior related one although the evidence bases obviously are kind of one and the same.

So, one, it was old. I had a little bit difficulty understanding the site that's 41 percent gap in 2008 and then gave percentiles on – I had a little bit difficult time interpreting that. But super boarding all of that is the fact that it's old. And I thought that the disparities thing, as Mahesh mentioned, is talking about access but that's not what we're talking about. We're talking about adequacy.

Poonam Bal: OK. And then anyone want to make a comment in reference to those comments or additional comments?

Peter Crooks: Well, digging into disparities a little bit more, they say that there was 40 percent adequacy rate in African-Americans in 1994 which was improved to 80 percent in 1997. And similarly, and I say 10 percent – Well, I'm kind of confused by this too. But apparently the gap in African-Americans and male, female has closed quite a bit. So, it's very important that, you know, that could be a no performance gap overall but there could be disparities that we're to be looking for. There is some disparities based on their data but those gaps seemed to be closing as well.

Poonam Bal: Let's open it up and see. Developers, do you guys happen to have any newer data on gap? I know you've reported 2008 data. Do you have anything more recent than that?

Female: Is anyone on from the Renal Physicians Association?

(Bane): Yes, hi. It's (Inaudible) (Bane) for RPA. And I don't have newer information in terms of the gap. What we have here admittedly is relatively old, a 2008 data, but expect that this parallel, to great extent, the facility data, although, as

Mahesh pointed out earlier, it is possible because it is physician level data that the numbers could be slightly different.

Michael Fischer: Let me ask a question about process. Assuming you have two measures such as this in the prior one discuss that will be a one facility that we wanted the clinician level. You judge the application on what's presented within it, in other words, we can't borrow from the evidence space with the other application provided to support this one or we are allowed to do that? I guess that's just – I'm just trying to make sure I understand how they properly interpret in a given measure application.

Peter Crooks: This is Peter …

Poonam Bal: Right. Well, so – Go ahead, Peter.

Peter Crooks. Well, I was going to say that the committee is free to use not only what the developer is represented but our own knowledge when it comes to – from the end.

Michael Fischer: OK.

Peter Crooks: And if they – for instance if they leave out something and we know about it or there's a new review or something, it's certainly our prerogative to take that into account.

Michael Fischer: OK. Did that help? I just want to make sure I understood. OK, thank you.

Mahesh Krishnan: OK, we move to Section 2 then for the interest of time. So for Section 2, we found the specification when I know we'll address this later on in the synchronization. It just – this happen – I happen to review the 0249 measure and then this measure. And so, the data definitions are not the same both in terms of the level of detail that's been provided as well as sort of the denominator of exclusion criteria, so something to think about.

I thought reliability testing was done and it was sufficiently reliable from implementation and performing gaps interpretation. And validity testing was done with a relative small number of physicians and patients but I still think it

was sufficient I think to be used and there were no threats validity that I know that – the risk adjustments that were provided.

And again, with regards to a meaningful differentiation, this again relied on 2008 data. I wish it was all that more recent data was not provided. So, looking at 2008 data, it looked like there was a need for meaningful differentiation.

But again, I go back to the prior question as that's what the case 17 years or 7 years later, there were no missing sets of data. It was doing multiple sets of measures and it's missing data. It wasn't an inclusion or exclusion criteria specified, so it's hard to understand the relative impact of that and it was not a composite performance metric.

Michael, you wish to add to that?

Michael Fischer: Yes, I – I thought that there's – this overall there was some parts of the – about the reliability and validity testing hard to read as – I think the faith validity based on the expert panel is not in question. I thought that there was some issues as you pointed out with specifications and needing to revisit that with greater detail. But when I read the validity and reliability I thought there were things that we're kind of repeated throughout the application which further may be bit difficult to read.

And then the other thing, as you pointed that, a lot of this is from 2008 and it's a small sample. If I remember correctly, it's couple hundred patients if it's right from the Midwest or maybe …

Mahesh Krishnan: Yes.

Michael Fischer: Yes. But it's relatively small and old. Anyway, I think that's all I have to add.

Peter Crooks: This is Peter. Under – just like kind of comment about their methods. When you test critical data elements and I'm not statistician or – but from discussions with the staff and also from our proposal, when you do critical elements that can actually be use for reliability and validity. And if that's

case, they'd miss the opportunity to mention that they'd already tested critical data elements in the validity section.

So – And this is going to be, you know, kind of recurring thing overlooking at reliability and validity if they actually go and verify critical data elements that also speaks to validity.

Poonam Bal: So, this is NQF staff. Let me make sure that everybody is aware, you're almost right except for one little thing. It depends on exactly their methodology that they looked at with their data elements.

So, if they compared, for example, something from a claim to the actual medical record and look to see that it was correct, then that would be validity as well. But if they just looked at say two different folks pulling data from our record and compared what – the two different abstractors found, that would be considered data element reliability but that wouldn't count for data element validity.

So, does that make sense?

Peter Crooks: Well, thank you for – Yes, thank you for that clarification. That makes perfect sense.

Poonam Bal: OK. And I couldn't tell. I assume since they talked about using inter-rater reliability, I assumed that they were comparing results of two abstractors.

Peter Crooks: That's exactly what they say. So, the critical data element testing reliability was just with those inter-rater reliability not a secondary source for checking the validity of the element.

Poonam Bal: Right.

Peter Crooks: So, thank you.

Mahesh Krishnan: Great. All right, let me go through sections three and four and five all at once then. So, feasibility, in my opinion, the data elements that are required for this

are routinely measured as part of the optimal care until the collection of this data is feasible and transmittable.

From a usability and use perspective, the measure will be used for physician performance and the accountability application for which they're being used is PQRS physician compare and the RPA internal quality improvement initiatives. So, that made sense to me.

And lastly, with regards with to competing measures as we measure – as we said before, while it's not specifically a competing measure and it's measuring physicians, I do believe that harmonizing this measure with the clinic measure in terms of the data definitions for 0249 makes a lot of sense to me. If not, it could be an apples and oranges comparison given that the same patients are being evaluated with two different definitions for two different stakeholders, the clinic or the physician.

Michael Fischer: A question for the staff. Was this measure – Were these two measures harmonized in the last go around so that the specs were the same?

Peter Crooks: I think that happened for the P.D. but maybe not for a hemo.

Mahesh Krishnan: I think the note said in the thing that was done for P.D. I don't remember saying that hemo, Peter.

Peter Crooks: OK. Yes, it certainly seems that there's an opportunity for that, for these two measures.

Poonam B al: Yes, definitely. And we'll be discussing that further in-person meeting.

Peter Crooks: OK. Thank you.

Poonam Bal: Well, thank you so much for that quick analysis. And again, we will be going into more detail and going step by step in-person meeting but that was a great synopsis.

So, we could move on to the next measure which is 0308. And that is – that was with led by Lisa and Josh. Unfortunately, they were both unable to attend

today's meeting. Is there anyone that would like to volunteer to take the lead on this measure?

Michael Fischer: Looking through my notes. Which one is …

(Crosstalk)

Mahesh Krishnan: It's relatively similar. It's OK. This is the parallel measure to hemodialysis adequacy for facility level, 0318, delivered dose of P.D. above minimum. The description of the measure is the percentage of all patient months for which patients – it says equal to 18. I assume greater than or equal to 18 whose delivered peritoneal dialysis dose was Kt/V between – of single-pool of Kt/V of 1.7 and the single-pool of Kt/V is less than 8.5 (dialytic) plus residual.

And the rationale was similar to the rationale for H.D. adequacy that P.D. adequacy every four months and the dose is critical to – in short time dose adjustments, 1.7 for adults, 1.8 for pediatrics, and as relates to patient outcomes, therefore, continue the implementation of this measure as reasonable that – the numerator of this is number of patients – a number of patients in the denominator who's delivered peritoneal dose was a weekly Kt/V between 1.7 and 8.5 single-pool (dialytic) plus residual.

And in the denominator statement is to be included in the denominator for any particular month. The patient must have had ESRD for greater than 90 days, be greater than or equal to 18 months, and be assigned to that facility for one month. And the denominator exclusions were pediatric patients, I.E. less than 18 years of age, patients who had ERSD for less than 91 days and patients who've not been in the facility for an entire month.

So, maybe just to buzz through this really quickly. The evidence or gap was similar in terms of hemodialysis adequacy. The adequacy guidelines from 2006 or KDOQI were mentioned as well as some data from perspective RCTs. And I thought this was reasonable. The performance gap that was illustrated via CROWNWeb and Medicare claims data from January to December of 2013. And means results indicated on average, facilities were meeting this criteria were 79 percent of P.D. patients and the sample size was large. It was

45,500 patients at 1,500 facilities. And the developer presented data on disparities and they know there were no discrepancies on disparities.

So, I don't if I'll – I mean, I'll just pause there and see if people have any thoughts as I look into the next section quickly.

Michael Fischer: Yes, let's back up to the evidence for a second. I think it was notable that they didn't mention ADEMEX in the evidence review which doesn't change their goal I don't think but I think it's an important – obviously an important study.

And, also – but I guess that give some specs, OK. So, that's my one comment on the evidence.

Mahesh Krishnan: Yes. I think the other comment – I think that makes sense. The other comment I think was the same discussion around the relevancy of the upper round and where that came from which I don't know if, Joe, you'd be able to address, but that's – if anyone else would comment, but that's the same parallel questions we did in the hemo side.

Michael Fischer: Well, he's maybe coming on. The other thing is – and I guess it's a sort of combination specs and I don't – you know, the evidence says they should include residual kidney function in your calculation. And in the specs, I'm not sure that they do or I saw the question how they take the – is the residual function taken into account when they're actually doing the specifications and the measure.

Male: So, this is (Inaudible). I can comment when NQF staff say it's appropriate for me to do so.

Poonam Bal: Go ahead. Thank you.

Male: Yes. So, residual kidney function is part of the P.D. Kt/V metric consistent with the KDOQI guidelines and with the previously endorsed statements. So, residual renal function is there.

Regarding the higher limit or the higher bound, it's the same rationale as previously. It's clinically implausible to achieve a Kt/V of 8.5 in a peritoneal

dialysis patient without enough residual kidney function to render. They're not ESRD eligible. And we are bumping up against those two markers, 8.88, 9.99. So, that's the only justification.

Mahesh Krishnan: I have a question for the NQF staff actually related to Peter's questions and Joe's response.

As we think through some of these measures and the measure definitions, there are a lot of micro specifications which underlie the calculations, some of the nuances we talked about before in terms of, you know, seven treatments or including this or including that. Having tried to replicate through these measures myself and understanding this variability, to what extent do we want to see sufficient clarity in the measure – metric definitions that addresses some of those micro specifications at least for the data elements as opposed to some of the other nuances in micro specifications that may be relevant to how you process claims like 8.88 and the 9.99.

Could you guys address that for us?

Poonam Bal: Well …

(Off-Mike)

Poonam Bal: I just kind of think about it. Is it clear enough so that different folks could apply the specs and come up with similar answers if they have the same patients? So …

Mahesh Krishnan: Yes.

Poonam Bal: OK. I don't know that I can get a better answer than that for you but it …

Mahesh Krishnan: No, that's …

Poonam Bal: Yes.

Mahesh Krishnan: Yes, that's perfect. I mean, that would necessitate a certain level of detail, right, because it's not, to your point, sometimes people make different

assumptions and then even though they say they're calling a banana, sometimes it's round and it's orange, right?

Poonam Bal: Right.

Peter Crooks: Yes, this is Peter. I would like to echo your concern that, you know, that there's a numerator brief statement and there's the numerator details, and I've been impressed on many of these that the numerator details are really lacking. You can't figure out from that how they actually determined it I think a little more detail. But also they have data dictionaries are available. In some cases, I've got the data dictionaries and I haven't always been able to find what I'm looking for.

So, I guess my general comment is that I think developers could be putting a little more meat into the detailed numerator and denominator descriptions.

Male: Yes, we could.

Mahesh Krishnan: And it's actually critically important when we do the CROWNWeb data transmission because we essentially take data elements out of an EMR try to match them to what we think people are asking for and then we submit them as such. If there's not sufficient detail, sometimes you have to make assumptions as to what's in there and sometimes people make different assumptions.

So, I think the data – and we are going through an exercise with CMS to go through some of the data definitions or having that degree of specificity in numerator, denominator exclusion I think would be extremely helpful.

Michael Fischer: Yes, this is Michael. I mean, on that theme and not to be too basic, if you read the denominator statement, anyone with ESRD is included. It's not even – so, hemo patients, if you read that denominator statement, would be in the denominator. Or perhaps, I have a misunderstanding.

Peter Crooks: Yes, that's a little bit loose, yes, right …

Michael Fischer: I mean, I'm not trying to – I get that we all realize that's not the case, but just touching on the theme of lack of specificity, that denominator statement and the denominator exclusion leaves chronic hemo patients in the denominator.

Peter Crooks: Right, right. Oops.

Poonam Bal: So, the developers that are on the line, there's going to be time between now and your in-person meeting to clarify some of these things, possibly not right now in the forms because it's kind of confusing to go back and add to the forms but maybe let's think about – and maybe you can talk with the project team a little bit later, think about maybe some additional stuff that you can supply and then kind of on the backend, we can get those into the forms for you somehow. But we can talk about that later.

Mahesh Krishnan: OK. Let me keep moving through these relatively quickly.

OK. Specifications. I think we just talked about that. So, specifications requires to be consistent and reliable. So, the specifications were given in here. I think, as I just said, the only comment I had were the specificity of what we just talked about. So, I'm not going to be redundant and cross over that again.

And reliability …

Peter Crooks: This is a – I've got a question about specification. Another …

Mahesh Krishnan: Go ahead. Yes, yes.

Peter Crooks: So, the – the numerator statement says once in four months. And so, I'm not clear how that's done. Is it that – And again, this is where the lack of detail how this is done, but I guess you have a patient in the denominator, then you look up to see if there's a Kt/V at – or set some dates, three months prior to the date of the measure I'm guessing, but is that sort of the approach or is there a code that someone puts in and says, "Yes, it's been done within four months," and we're going to just put that code in? How is that actually done? And either the developer can – is a best to answer that.

Poonam Bal: Yes. Developer, would you like to respond?

Anyone from the Renal Physicians Association on? I'm sorry.

(Crosstalk)

Male: That's a CMS measure.

Joseph Messana: Right. So, this is Joe Messana and I'm at a tech meeting in Baltimore and I don't have the information in front of me. And so, I would respectfully request that we defer to that question till the in-person since we're providing additional detail for the working group. I would suspect that we should have the opportunity to provide that additional detail for Dr. Crooks and others at that point.

Peter Crooks: That's fine.

Poonam Bal: That's …

Joseph Messana: OK. I'm in a hotel lobby and I'm just not in the position to answer that right now. I'd rather not make something up for you.

Peter Crooks: Yes. We appreciate that. That'd be fine.

Joseph Messana: OK. All right, thanks.

Mahesh Krishnan: OK. It's time that we get to move to reliability testing.

Michael Fischer: I think so, yes.

Mahesh Krishnan: So, reliability testing was done using CROWNWeb and Medicare data, I suspect using the same formula that was previously discussed. Validity testing was done as well. And the evidence providers suggest that evaluation of P.D. be conducted every four months to ensure timely dialysis dosing and adequate dose. And then for validity testing itself, it was calculated using (inaudible) here, using the experiment correlation between this specific measure and the 2013 standardized mortality ratio and (standardized) hospitalization ratio.

It was also being – The measure is also being maintained on the basis of faith validity and the correlation between this measure and the 2013 mortality ratio was negative 0.08 with a P value of 0.7744, and the correlation between the measure and the hospitalization ratio was negative 0.139 with a P value of less than 0.0001. And there was a correlation which associated higher facility level percentages of this metric at a facility level to achieve with lower standard hospitalization ratio although the magnitude of the association was low, a weak association between facility level percentages of patients receiving with adequacy measure and standardized mortality was observed in the expected direction, however, correlation was not statistically significant.

So, if I just go to this section two completely, there were no measure exclusions, there were no risk adjustments.

In terms of meaningful differences, differences in the measure were evaluated separately using patient level values for achievement versus non-achievement. But the performance, it's to show that 1,270 facilities which was about 83 percent achieved the expected performance and 258 facilities, around 17 percent, performed worse than expected.

In general, there was discussion around the fact that higher Kt/Vs represent higher quality of care. And so they felt that that was reasonable for meaningful differences.

Patients with missing data event, patients with missing data were not excluded from the measure. Therefore, patients for home Kt/V value was missing for the month or still unclear in the denominator and this was deliberately – this was designed to ensure the facilities will still be evaluated for the measure.

I didn't quite follow that when I read that. So, I'd love an explanation of that.

And then the last section before we go to general comments was the composite measure. And there was not a composite measure in there. And there were no specific pre-evaluation comments for this.

So, that was my only question on this was I didn't understand that section around patient with missing values were not excluded from the measure. I just didn't understand that.

Peter Crooks: This is Peter. My interpretation to that is you keep them in the denominator and then if they don't have data, it says if they were negative, you know, because they don't get down to the numerator, which is I think is the right way to do it. Is that what you were wondering about or is that not answering your concern?

Mahesh Krishnan: Yes. Was that the intent from the developer?

Peter Crooks: Yes, the intent is to – yes – is to not let gaming go on where – so, we don't have a data so that patient is not accounted. They're still accounted – it's accounted negative.

Mahesh Krishnan: OK. Yes, got it. Makes sense.

Peter Crooks: My comment is just briefly I think it passes on reliability, validity, and so on.

Mahesh Krishnan: Yes.

OK. Any other questions or comments on that one?

All right. Let me brief through – Sorry, go ahead, Peter.

Peter Crooks: Well, feasibility and usability are clearly meeting the criteria.

Mahesh Krishnan: Yes, exactly. I think all of the caveats around CROWNWeb data that we talked about before for the hemo measure probably apply to this as well. But the rest of that, usability and – it seems not. And then for the linear related to our competing measures when we talk about the P.D. – the parallel P.D. metric that was relating your competing measure, that should be harmonized.

Michael Fischer: May I ask another question. This is Michael. You know, when I was reading the evidence for the measure, it says that adequacy should be measured in the languages at least every four months. And if you read the measure, that's the evidence that that is presented. And if you read the measure, you know, it's

written as number of months. So, it seems like there's an implicit presumption that it's being measured every month.

And I guess I just – when I was – went through again, is that a discrepancy that in terms of what the evidence is recommending and what the measure is or perhaps I'm focused on that inappropriately. I just was curious what others thought.

Peter Crooks: This goes to my question about the specifications too that they and – I think they'll clarify for us how exactly they do that. I think what they're really requiring is that is done within the last four months of the date of testing but it's not clear how that's done, whether there's some code that people put in or they go and look at the lab date or …

Michael Fischer: OK. So, thanks, Peter.

Mahesh Krishnan: I mean, in CROWNWeb, the data elements are done by date. And so, presumably to your point, Peter, you would just – you would look at it within beyond of a X number of months and see if there was a value within that.

But maybe, Joe, you could comment if that's what you guys were thinking about.

Joseph Messana: Yes. This is Messana. So, in fact that's what the measure specified. What I'm just not certain about and I'm not prepared to discuss is exactly how we define the denominator of that four-month period. But the intent is consistent with the KDOQI guidelines and the calculations of reliability and validity are reflective of that every four months. So, we can provide you a bit more detail but I will reassure you that it's one Kt/V or the number of patients who have a Kt/V in a four-month period. And we'll provide a little bit more detail about the calculation logic for the in-person meeting.

Mahesh Krishnan: Yes. And the CROWNWeb versus claims, yes, that'd be great.

Poonam Bal: OK. Were there any additional comments before we move on the next measure?

OK, thank you so much for that.

And so the next measure we want to discuss is 0321 which is Peter and Mahesh. And this is also a Renal Physicians Association measure. So, if you want to start with the quick overview and then any highlight that you want to bring up.

Peter Crooks: Yes, Mahesh would be first – go ahead.

Mahesh Krishnan: Yes. Peter designated me to designate it better on this one. So, this one is the companion measure to the one we just talked about. It's from the RPA, it's for physicians.

A brief description of the measure, percentage of patients aged 18 years or older with the diagnosis of ESRD receiving peritoneal dialysis who've had a total Kt/V of greater than or equal to 1.7 per week measured once every four months. The rationale was that the dialysis dosing is strongly associated with better outcomes including decreased mortality, hospitalizations, fewer days in the hospital.

The calculation was, in the numerator, patients have a total Kt/V of greater than or equal to 1.7 per week and measured once every four months. Denominator is all patients grade – age 18 or older with the diagnosis of ESRD receiving peritoneal dialysis and there were no denominator exceptions for this measure.

To go through then the commentary on this one. For evidence support, in my opinion, there was sufficient evidence to support this similar to what we just talked about before. There was – I was going to section one all at once. The performance gap was again based on older data and they are 76.5, 8 percent of patients were reported to not have received optimal care which again is a very old data.

Similarly, a similar measure, the measure we just evaluated showed that this gap was about the same, 79 percent of patients met this measure utilizing more recent data. So, I believe that there is still a gap that one could look at.

The data submitted for disparities here indicated that there may be a gap to access to P.D. But I didn't see data supporting a differential performance by race. But in prior announces for the clinic level measure, there did not seemed to be disparity. Again, I thought it addressed a high priority topic, dialysis adequacy, the fundamental measure for physician prescription as well as dialysis clinic performance. And there was no composite measure.

So, I'll just pause and see, Peter, would you add to that?

Peter Crooks: So, try to get my notes correct here. The – I'm with you so far, go ahead. I'll jump in later if I have something else to add.

Mahesh Krishnan: OK. Any other thoughts, guys? I will take the deafening silence as permission to move on to section two.

(Crosstalk)

Peter Crooks: … that they have the same question. I wanted to make sure that they're including residual volume in their metric, and as suggest in the evidence review, that should be part of the Kt/V measurement.

Mahesh Krishnan: Yes. I think that speaks to the concerns I have in the specifications section for section …

Peter Crooks: Right.

Mahesh Krishnan: … two which again, you know, I just think that we need better, more detailed definitions here, and again, the harmonization between the two measures from the specifications which I know has happened before but it's no longer the case would be very helpful.

For reliability testing perspective, I thought the results just demonstrated sufficient reliability to be used for best in performance gap identification. For validity, again, it was this relatively small number of physicians and patients, but I thought it was sufficient that could be used. I saw no threats to validity. I saw (inaudible) risk adjustments.

Again, for meaningful differentiation, 2008 data was provided. And from there, it seemed like there was sufficient data provided to demonstrate (inaudible) for meaningful differentiation but more recent data would have been helpful. There are no multiple sets of measures.

The missing data question wasn't really addressed specifically because in the data definition and it was not a composite metric definition for this.

Peter, your thoughts?

Peter Crooks: Well, you move fast. So, I'm trying to go back to the, yes, reliability and validity. They did inter-rater reliability clinical data element and that looked OK. For validity, they more or less relied on a faith validity, the expert panel. And that was – they didn't try to do any other validity. They didn't try to correlate it to the outcome, to mortality or other things. So, you know, it's – the committee has to decide if that's – if they agree that that's valid enough just the expert panel.

They do try to make a case that this – they can show statistical differences and meaningful differences. I'm not sure that that's really shown by the data that they give that they can actually sort out one group from another.

So, those are my concerns regarding, you know, the scientific properties of the measure.

Feasibility, it looks good and it is in current use. So, I think those are OK.

So, we'll open up for other comments on this measure.

Mahesh Krishnan: And can we just ask (Steve) (Inaudible), I assume that's the same issues here, right, where we don't have more recent data than the 2008 data.

(Steve): Yes, that's right. Not that I'm familiar with. By the time of the in-person meeting, you know, I'll check internally in the organization to see if we do have further testing, but I'm guessing not.

Mahesh Krishnan: OK.

Poonam Bal: And just to clarify, we don't necessarily have to have updated testing but if you have data from your – I'm sorry, I'm using the name of it, your quality improvement initiatives or anything like that it just give some flavor of where the rates are, where the performance rates are, that would be very helpful for the committee.

Mahesh Krishnan: Great.

Poonam Bal: We'll move forward.

Peter Crooks: Into the developers' credit, they did use ADEMEX in this evidence review that isn't – they didn't the CMS for the record. OK.

Mahesh Krishnan: OK. So, maybe just to cover sections three, four and five at once, feasibility, usability and use and competing measures.

Feasibility, again, as mentioned before, the data elements require routinely measured as part of patient care. So, I think collection of this data is admittedly feasible.

For usability and use, again, this would be used for physician performance, accountability application is for PQRS, physician compare and the RPA internal quality improvement initiatives.

And then lastly, with regards to competing measures, I do think that this measure had previously been harmonized with 0318, but now appears to be out of sync with 0318. And so, my – I would advocate as I did for the hemodialysis physician measure that we harmonize these measures in terms of specificity of the data specification as well as the operating definitions.

Peter Crooks: Yes, Mahesh, and just to – for process check, we're not going to be considering harmonization of related and competing measures on the workgroup calls. Not in here.

Poonam Bal: Yes, that's correct. We will go over that in the in-person meeting but we generally tend the hold the discussion off until all measures that we consider related are competing have been discussed and they've received your yes or no

for endorsement separately. And then once that's done, we discuss them as a group. Any measures that have been recommended for endorsement, we would discuss those.

Mahesh Krishnan: Got it. Yes, makes sense, and you know which ones are harmonized, which one too.

Peter Crooks: Right. If something is not going to be endorsed, we don't want to take time trying to harmonize it.

Mahesh Krishnan: Yes, yes. Makes sense.

Peter Crooks: Thank you. Great. Great review. Thank you.

So, other people have comments on – they'd like to make, any part of this measure, 0321?

Poonam Bal: OK.

Peter Crooks: Great. Well, we have a chance to actually getting through these all today. This is great.

Poonam Bal: Yes. The committee is doing a great job of getting through the measures and still have an internal discussion. So, we really appreciate all the hard work that everyone's put into reviewing these measures beforehand.

So, I will move forward to 0255. Michael Fischer and Ishir Bhan are on.

Michael Fischer: Yes.

Ishir Bhan: Want to take this? Either way is fine with me.

Michael Fischer: I'm happy to walk through it, Ishir, or if you want to, that's fine.

Ishir Bhan: Why don't I start. I might unfortunately drop off the call in a little bit, and I could start to that.

So, this is a – just starting with the beginning here. A brief description here is this is a percentage of all P.D. and H.D. patient months with serum or plasma

phosphorous measure to at least once within the month. And the rationale here is consistent moderate in phosphorous level, helps ensure the regulation of patient morbidity and mortality including (inaudible) of bone density decrease on pain. Fracture prevention and decrease rate are (inaudible) related condition. E.G. stroke and heart attack and routine blood test for also aiding the detection and monitor of and monitoring for abnormal states (inaudible) of here. Phosphorus balance in this, especially going on population.

The numerator is the number of patients and also patient months in the denominator with serum or plasma phosphorous measure at least once within the reporting month. Denominator is number of patient months from an in-center hemodialysis, home hemodialysis or peritoneal dialysis patients under the care of the dialysis or the – or with – for the entire reporting month.

(Off-Mike)

Ishir Bhan: And, let's see. Exclusions are implicit in the denominator. Definition include all patients who have not been in the – for the entire reporting month. There are no additional exclusions for this measure.

OK. So, just hopping over to evidence. There's a process measure and the developers provides the following path between the measure focus and the health outcome measure serum or plasma phosphorus assessed value, identify the problem, identify treatment options, then administer the appropriate treatment. And after going through all those steps, hopefully the patient experience it increases then mortality.

They reference the KDIGO, clinical practice guidelines for CKD-MBD and the specific citations here are that in patients with the CKD stage three to 5D, the – it's a reasonable base of frequency in monitoring serum, calcium, phosphorus and PTH on the presence of biochemical abnormalities and rate the progression of CKD. And they say in stage five, CKD including 5D, calcium and phosphorus should be measured every one to three months, and for PTH every three to six months. In CKD patients receiving treatments for CKD-MBD or in the biochemical abnormalities are identified, as a result, to

increase the frequency to monitor for trend in treatment efficacy and side effects.

So, questions for the committee are, is evidence directly applicable to the process of care being measured, is the process of care proximal and closely related to the desired outcomes.

And the sort of commentary, are there or could there be performance measures to related health outcome more evidence-based, intermediate clinical outcomes, intervention for treatment, and is there evidence of systematic assessment or export opinion beyond those involved in developing measure and does the (inaudible) agree that it is accessible or beneficial to most providers council without empiric evidence.

So, I'll just start out by saying that there is quite a separation here between the process of actually measuring the phosphorus and then down the road on the clinical improvement, but of course that is – if you don't measure the phosphorus, you're not going to be able to manage it.

There is reasonable wealth of information around phosphorus levels and outcomes with higher phosphorus levels being linked to deleterious outcomes in ESRD, although it's not cited in the detail of the manner here. But again, there's a lot of steps pertaining to measurement and affecting change, but let me open up the statement – the floor here.

Michael Fischer: Well, this is Michael. So, I mean, there's no – there's growing observational data, as you point out, about phosphorus levels and outcomes. But this is about the act of measuring the phosphorus. So, I think, you know, there are steps between those two things in terms of what's done after you measure phosphorus.

I guess it's important to remember that this is a measure that was endorsed in the past. The evidence base was felt to be sufficient. I realize we're evaluating it right now. But there's only been a growth in observational data in ESRD about high phosphorus levels and worse outcome. Again, that's not the same as what this is addressing which is actually just measuring it.

Mahesh Krishnan: I think it's a routine process of care marker, right, just as something that should be done in the dialysis facility, it makes sense to me. I got a little bit lost in the story line between the connection, lots of data – retrospective data between the connection of serum phosphorus and outcomes, limited data except for what was present here in terms of correlation later on between the measurement as a process marker and outcomes.

But that being said, I just feel a reason to play confident in thinking through that it's a reasonable process marker. I don't about the exact link to mortality and morbidity, but I think it's reasonable to look at.

Ishir Bhan: OK, any other comments on this?

Peter Crooks: This is Peter. I'd like to say that in past reviews, this has always been a difficult one. We are – We have to make a leap of faith that – to that correcting phosphorus actually leads to better outcomes. This is, you know, as we all are aware, there's a lack of clinical trials to demonstrate that. There's a wealth of observational studies showing that high phosphorus is linked to higher mortality classification of arteries, and so on, but it is leap of faith for us re-endorse this again if we do. That's not only measuring but actually knowing the values and then acting on that will impact mortality. We really don't know.

This one barely squeak by in the last time or two, as I recall, because of evidence.

Ishir Bhan: So, why don't we – just moving on to the performance desk.

Poonam Bal: And this is NQF staff. I just want to make sure that everybody is familiar with the bottom part of the evidence algorithm if you go through. Just a reminder that expert opinion is not considered evidence by NQF but if you keep going down in the algorithm, there is some questions there to guide you through thinking about a potential exception to this evidence criteria. And that's why these questions were actually asked to you for you to consider.

Mahesh Krishnan: I'd be curious though, could you mean the prior statement, the guidelines are considered a higher grade of evidence though sometimes guidelines, when

you tie into them at times are based on data and at times are based on expert opinion. How do you reconcile those two?

Poonam Bal: Right. Well, like you said, different guidelines have different levels of depth, if you will. So, some guidelines are very much evidence-based and you would just take, you know, you decide what they have. Some recommendations are based on evidence and you would use those. Others they say right out that it's just expert opinion. So, even if there's a guideline, just because there's a guideline doesn't mean that it goes all the way through. You have to see what the level of evidence is that goes with it.

Michael Fischer: This is Michael. And with this particular one, if I'm correct, KDOQI recommends it, but there's – without any grading.

Peter Crooks: Right, right. Which are …

Poonam Bal: Right.

Peter Crooks: … this is kind of muddiest picture. Following through this algorithm, it might get to number six, does the grade for evidence recommendation indicate high quality evidence versus no grading. And if we follow that, would we have to rate this evidence as well if it's even acceptable.

Poonam Bal: The other way to do it, and I don't – I haven't looked at the details of the measure, but even if there's not a grade because not every guideline does a grade, if you can pull from the information that the developer gave you, the quantity, quality and consistency, you may be able to figure out from that where to land. So, really, you kind of have a couple of choices if they give you QTC, you can use five A, B, and C to get you where you need to be. If they don't give you that but you know the grade, then you can use the grade. Otherwise, you're probably a little bit lower in that algorithm.

Female: And we definitely recommend that the committee use the algorithm during the workgroup calls and in-person meeting to assist you in making your decisions.

So, if no one had any other comments, we can move forward to gap.

Ishir Bhan: OK. So, for this performance gap, performance scores were using within January 2013 CROWNWeb clinical data from 6,073 facilities having at least one eligible patient based on the patient month with a mean and standard dev of 87 percent to 18 percent respectively, a minimum of 0 percent and the maximum of 100 percent. 25th percentile was 86 percent, 50th percentile was at 92 percent, 75th percentile was 96 percent, and the developer did provide disparity analysis performed along the entire eligible population. 518,127, the results are considered to imply statistically significant changes in performance score, according to (inaudible), et cetera. And the developer did not feel these were clinically meaningful.

So, the question for us, so there's questions on the gap and care, is there a gap and care that warrant the national performance measure and should this measure be indicated as disparity sensitive.

You know, I was kind of struck by these numbers. They're a little lower than I expected. Although even that said, with the median of 92 percent, they're fairly high overall. But measuring phosphorus, you know, they're pretty automatic trigger at least in the facility type theme.

Mahesh Krishnan: Yes. So, I suspect the difference between what you perceive and what you see in the data is twofold. One was the data transmission issue with CROWNWeb, right, because there's no claims backup here. So …

Ishir Bhan: Right.

Mahesh Krishnan: … that's just – it is what it is and, in a nutshell, if you submit 100 patients today because the matching algorithms are slightly off, only 90 of them may make it through. So, that becomes an issue. And then two, there was a period of time which includes 2000 – the calendar year 2013 where the facilities where instructed to auto discharge patients for extended hospital fees, and so that also screwed up to denominators just for a good while which is why I'm concerned about to submit discharge stuff that continues go on.

Ishir Bhan: So, again, there's this question of, you know, we're starting with a median of 92 percent. Is there really much room from improvement here, especially given the caveat that you just mentioned of the data may not reflect the reality

that (inaudible) missing because it's based on CROWNWeb. And, you know, further emphasizing the nature that this is quite – just separate from the actual clinical outcomes, the measurement of phosphorus.

I know there's some skepticism in my voice, as you my get.

Mahesh Krishnan: Yes, I agree.

Peter Crooks: OK.

Poonam Bal: And definitely you bring these topics up again during the in-person meeting and the group can discuss it further. And we need to determine if we want to move forward beyond gap.

So, at this time, though, we will move forward to reliability and validity, keep the discussion going. Just a reminder, in about five minutes, we'll need to take – if we don't finish all those, I doubt we'll be able get to this measure and the next one on the next five minutes, we need to pause for public member commenting. But please continue discussion until then.

Ishir Bhan: All right. So, why don't I keep cruising then so we can make it through this? Reliability, the process measure – it's a process measure associated in dialysis patient and had their phosphorus level measured at least once during the month measuring this CROWNWeb data and the field from abstraction and from CROWNWeb are defined. And staff were unable to identify how dialysis based are identified and the measure is not risk adjusted or stratified.

So, we'll see the questions here, but I think clearly the definitions are pretty straightforward. It's really a question of the data itself as far as I'm concern, but I'll let people chime in.

Mahesh Krishnan: I agree.

Peter Crooks: Yes, keep going.

Ishir Bhan: All right. So, reliability testing. Testing was conducted at the facility level. It's in CROWNWeb 2013 from nearly 6,000 facilities who are presenting both adults and pediatric patient performance measure score with the data source

CROWNWeb. Reliability was assessed by calculating at the facility level peers and correlation, coefficients between the current performance month of the preceding months and interim unit, reliability is also assessed and the coefficients were 0.72 to 0.90 and it's statistically significant. The monthly IURs rated from 0.95 to 0.97 and testing demonstrates that most variation can be attributed to between facility variation and peers and correlation indicated reliability over time.

So, any comments here?

Mahesh Krishnan: Nope.

Ishir Bhan: Seem pretty reasonable, OK. So, let me …

Peter Crooks: OK.

Ishir Bhan: OK. Excellent.

Validity evidence demonstrate importance of frequent monitoring in phosphorus level to identify needs for treatment adjustment and monitor morbidity guidelines and are overly prescriptive on the frequency of phosphorus level measurement are these specifications (inaudible) evidence. So, again, here, the only question was whether actually just measuring the phosphorus level and all that valuable here in and of itself sort of necessary but not sufficient.

Peter Crooks: Yes, they do show a correlation with mortality of doing the measurement with mortality from their database to – for the validity test.

Ishir Bhan: Yes, it's just we can – I just quickly run through that. So, the developer utilized CROWNWeb. And you can put on regression models basically came up with the relative risk of mortality for 10 percent increase. The performance score is 0.98 so that's a two percent reduction in mortality statistically significant that's where a 10 percent increase in the performance score and my only concern here was restarting with the median of 92 percent. So, there's, again, a 0.02 reduction in mortality that really would be talking about a small

number of facilities, so that could be effected at greater magnitude to be affected.

Mahesh Krishnan: Yes, and just to add additional data, we've done this interim analysis at DaVita and we've essentially frozen the phosphate levels where they are because we don't think meaningful improvement is – makes more sense, and to your point, there are additional measures that we could concentrate on above the fundamentals where we actually showed improvement we can actually improve mortality and morbidity. So, our data supports that as well.

Ishir Bhan: OK. There were no exclusions. And …

Mahesh Krishnan: I have a quick question on exclusions though.

Ishir Bhan: Yes, yes.

Mahesh Bhan: If a patient is hospitalize – OK this is the question I think it was raised someone in there if a patient is hospitalized because like the minimum number of treatments the patient experience so if a patient is hospitalize and we obtain the phosphorus and we type in back in to the computer system that means it's necessarily reflective for the patient level of care but I just wonder whether it is going to back to whether not patient to don't have a minimum number of treatments in the facility ought to be included or not. I think that would actually hone the definition a little bit more.

Ishir Bhan: Yes. It's not clear, yes.

Poonam Bal: And since that we are running out of time, so I will ask the developers to take that feedback and be prepared to respond to that into in-person meeting and I know we're in between the measure now but I do want to pause a discussion just to see if there's any public member comments and then we'll go finish this measure off and do a really quick review of 1454 before the end of the day.

So, (Bridgette), if you could open up for member and public commenting, please?

Operator: And at this time, if you would like to make a comment, please press star then the number one on your telephone keypad.

And there are no comments at this time.

Poonam Bal: OK, perfect. And then we'll continue the discussion for 0255. But let's go ahead and forward to feasibility and usability and so we can try to go a little bit of discussion in 1454.

Ishir Bhan: OK, so feasibility here. The data source with CROWNWeb as we talked about the pros and cons with that 5,951 facilities that at least 11 knowledgeable adults and pediatric patients during this period were included in the analysis and data are collected and generated during – and used by health care during a routine provision of care.

Yes, this is routinely generated. Looks like the data was readily available but the question was to whether is complete or not.

And just why don't I quickly also wrap up the usability and use, used in ERSD QIP, publicly reported in QIP year 2014. And public – the measure in – for ESRD QIP (inaudible) for at least eligible patients for the measure. So it is (both) reported. And I just open up for any closing comments and we can move on to the other measure.

Mahesh Krishnan: I'm just wondering. I don't remember this clearly but maybe we can address in the live session. When this measure was included in the past, I thought it was done attestation not by CROWNWeb. So I mean it's something to be nuanced even though the measure was done. I thought it was done slightly differently.

Peter Crooks: Good point.

Poonam Bal: OK. And we recommend that. We discuss that further in-person meeting. But I do want to get to this last measure just to give – well, we will probably won't have enough time to discuss early. Give it a couple of minutes. Unfortunate, this is another measure that Josh and Lisa was supposed to do

together. And they were both unable to attend today. Is there anyone that feel comfortable doing quick overview of 1454?

Peter Crooks: This is Peter. I can give a shot.

Poonam Bal: All right. Thank you so much.

Peter Crooks: Measure 1454, Proportion of patients with hypercalcemia measure steward is CMS. This is a new – or this is a re-endorsement too. I guess this was passed first in 2011. The brief description pertains of adult patients, with the 3-month rolling average of total uncorrected calcium greater than 12 – 10.2 numerator and denominator are pretty straightforward.

They do describe in the aspects when we get down to how they do this 3-month rolling average which data they use if there's more than one calcium in a month and so on.

The evidence is based on KDIGO 2009 and KDOQI 2003. The evidence for KDIGO is rated as 2D. I think it was rating they gave which – and KDOQI from 2003 is not rated. So it's expert recommendations.

Other sources, they – the literature search for TAP meeting in 2013 that apparently recommended this level of this measure. There's no clinical trails cited in the evidence. There's no meta-analysis. They do say evidence on calcium-based binders versus non- calcium-based phosphate binders which is rather indirect. They do say that increase calcium balances associated with poor survival. So I balanced the evidence – basis is weak. Just as a quick overview.

Let me stop there for others comment.

Mahesh Krishnan: Yes, I agree. And then the one nuance here that we've been struggling with, Peter, is even though they're maybe existing retrospect of evidence-based there is lab to lab variance between that. So I don't know how you deal with that mean when you try to reconcile the evidence space where – even the clinical trial that was done or even – in one data analysis, it was done in one

setting but now you have multiple labs contributing to the data sample. Just further arose the evidence-based for me.

Michael Fischer: This is Michael. I guess for me and I think as part of our discussion, (inaudible). This is a more of a – kind of about a patient safety issue with use of vitamin D analogs.

Peter Crooks: Right, right. And that was how we got it by member now, back in 2011 it was really consider the patient safety net.

Michael Fischer: Yes, that was my recollection, Peter.

Peter Crooks: Yes. Yes. Thank you. And when you look at the gap, you know, this is another one where the gap is really low. The mean percentage is 2.1 percent with the standard deviation of 3.2 percent which overlaps a zero. The 25the percentile 0.3 percent, 50th percentile, 1.2 percent, so that would be the median. 75th percentile is 2.9 percent.

So the gap is very small. And are we really going to be able to close that gap? Is that something that should have nationally endorsed measure? It's something for the committee to think about.

Michael Fischer: This is Michael. This is another one that I thought about, is it worth having this in a reserve status?

Peter Crooks: Right. And I think that's good so we can consider when we're in-person and – you know, we say that it – if it needs all the other levels for approval that we could consider that.

Mahesh Krishnan: Yes. When we get down to validity and reliability, I mean this is another one where – as we looking at our units there's discrepancies between the reported rating of a facility much in the EMR and because the data transmission issues. So I would agree with that – either non-endorsement or reserve.

Peter Crooks: Yes, maybe that speaks to the – when they did reliability testing, the entry unit variation was 61 to 66 percent. So in that most variation was entry unit but that's a lot of entry unit variation.

Mahesh Krishnan: Yes.

Peter Crooks: Which may go to the issue of different labs or different, you know …

Mahesh Krishnan: Yes, it's a combination – in my opinion, it's probably a combination of different labs as well as the combination of the data transmission success, right? So, we don't have – we don't do (inaudible) but you can compare the (basic) M.R. to that reported CROWNWeb data that comes back and there are some discrepancies. And to your point, become such a topped out metric, you lose one or two data elements, you immediately change your performance.

Peter Crooks: Right. And validity was based on phase validity – from a TEP committee and then for a technical expert panel. And also they were able to correlate from their database that more patients with calcium over 10.2 did link to higher mortality. So, they use those for their validity arguments.

Feasibility seems OK because we're hearing more and more about how much going to trust it, even our CROWNWeb issues but it is all electronically collect the data. It's currently – for usability it's currently in dialysis facility compare and they're planning – I guess it's been accepted into a plan ESRD QIP for 2016.

So I'll turn over for other comments.

Mahesh Krishnan: I think I'll give you my big issues.

Peter Crooks: Yes.

Poonam Bal: OK. So do we want to or just finish up with usability and use. We are at the top of the hour, so I do want to respect on time. Were there any comments that you want to make (inaudible) before we finish up?

I will take that as a no. So, thank you for everyone with their great work on our second workgroup call.

Moving forward we'll have two more workgroup group, one this Thursday, one Tuesday. Again, committee members and the public are free to join any of the workgroup call. And just listen in and so on. So, that's available to

you. We will be meeting in-person on May 6th and 7th which is not too far away from now.

Just a reminder, if you have not receive travel information or if you have not filled out your travel information, please do so as soon as possible. If you'd not received the information, you can e-mail the renalqualityforum.org e-mail and we will make sure that you receive the appropriate information.

If you have any difficulties getting the SharePoint, again, contact us by the in-person meeting we do hope that all committee members have reviewed all measures not just the measures in their workgroup and how to view the public and member comment that have come in.

So, I just want to stop and see if there's any other questions before we conclude the call.

Michael Fischer: This is Michael. Will there be any kind of summary of the discussion today?

Poonam Bal: We did not make summaries of these discussions.

Michael Fischer: OK.

Poonam Bal: But we do post the transcript on the website. Other than that, we do ask that lead discussant take the role of provide – basically stating what were said in the workgroup calls and start the discussion in that person with that statement.

Michael Fischer: OK.

Poonam Bal: Were there be any other questions?

Mahesh Krishnan: No. This is really good. Thank you. I've got to drop for another call, but this was great. Thank you. Mahesh.

Peter Crooks: Yes, thanks everyone.

Mahesh Krishnan: Great. Bye.

Poonam Bal: Have a nice day. Goodbye.

END