

## **NATIONAL QUALITY FORUM**

**Moderator: Lauralei Dorian**  
**February 13, 2012**  
**3:00 pm CT**

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

Lauralei Dorian: Good afternoon, everybody. Thank you for calling in today. This is Lauralei Dorian from the National Quality Forum. Thanks for your participation on the call and also for completing your preliminary survey if you've already done so.

I also have Karen Johnson with me in the room and Karen Pace from NQF on the phone. So I'm just going to get started by doing a brief roll call to see who we have with us from our steering committee. (Amelia), are you there? Tom Howe?

Thomas Howe: Yes, I am.

Lauralei Dorian: Okay, great. Suzanne?

Suzanne Heurtin-Roberts: I'm here.

Lauralei Dorian: Okay. (Mark)? Okay, (Julie)? Lorna?

Lorna Lynn: I'm here.

Lauralei Dorian: Great. And Bonnie?

Bonnie Wakefield: Here.

Lauralei Dorian: Wonderful. Okay, now do we have any representatives from CMF on this call yet?

Linda Klingensmith: Linda Klingensmith.

Lauralei Dorian: Linda did you say?

Linda Klingensmith: Yes.

Lauralei Dorian: Okay, great. And from PCPI?

Diedra Joseph: Yes, Diedra Joseph is on the call.

Lauralei Dorian: Okay, from AAD?

Alison Shippy: Alison Shippy is on the call and Oliver Wisco.

Lauralei Dorian: Great, thank you. And (CQA)? Okay, they might join us later, hopefully. So just a reminder that the call today is open to the public and we will be taking public comments towards the end of the call and also a reminder to please keep your phones on mute when you're not speaking because that can interfere with other people's lines.

Female: Also, I just wanted to let you know that (Acumen) is here. We're the developer for some of the CMF measures.

Lauralei Dorian: Okay, great. Thank you. So I'll hand the call over to Karen Johnson then.

Karen Johnson: Thanks, Lauralei. Before we start the call I'd like to just go through the process of what the call will try to accomplish. First of all, we're going to turn over, as you know, we asked several of you to be leads (discuss) for several of the measures, so we're pretty much going to turn the call over to you as the lead discussant and we'll ask you to summarize the measure and then summarize the ratings and rationale and then highlight some of the areas of concern that came up.

And you don't need to necessarily read out all the different comments that came through unless you just feel like you really need to do that, but really just a summary of the areas of concern or differences that came through on the preliminary evaluation.

After you finish your summary we'll ask other reviewers of the measure to make any additional comments that they care to and then we'll just open up the discussion to the workgroup members just for any discussion about the areas of concern.

And as you heard, we do have developers who are online, measure developers and at that point you may want to ask them any particular questions or ask them to clarify things, that sort of thing, so they are here in order to do that for you.

And the other thing to keep in mind is that we really only have about 1.5 hours, even this is a two-hour call. We have to leave time open at the end for public comments, if any, so we won't have an incredible amount of time for each measure.

So part of our role as the NQF staff will be, number one, to clarify when necessary any of the NQF criteria or guidance and then also to kind of keep an eye on the time and keep things

moving as we go through. And generally, just so you know, that is usually something that would be done by your committee co-chairs but, unfortunately, (Jerry) and (Don) are unavailable to participate on this call today, so I'm going to try to do that.

And I will say that if there are -- once we get through the measures, if there are burning issues that you still want to chat about, that sort of thing, the SharePoint site, we do have the discussion forum there and you can continue any discussion that you may want to about different -- the different measures using the SharePoint site.

So with that, I'd...

Lauralei Dorian: Sorry, it's Lauralei here. I did just want to also mention that we are screen sharing today, so you should be seeing an agenda up on your screen right now. If you're not, then we recommend that you X out of the program and try to log back in.

But if that still doesn't work you have all of the documents that we'll be going over. It's mostly the Excel spreadsheet with those preliminary ratings and then we might bring the measure evaluation form up on the screen as well. So you'll have access to those via SharePoint and we've emailed them to you in case you can't get the SharePoint program to work.

Karen Johnson: And just one more change to the agenda, Measure 0494, the medical (homes) system survey, we want to move that to the end of the call and we just wanted to remind you, as you already know, we were a little bit late in getting that measure information to you, so we don't really expect to have, and as a matter of fact we didn't have, on Saturday morning, we didn't have evaluations on that because of the tardiness there.

But hopefully we will have a little time to at least field a few questions about that measure even though you haven't done the preliminary evaluation.

So with that, I think I'm going to hand it off to Suzanne who is going to take ((inaudible)).

Suzanne Heurtin-Roberts: Okay, let's see. Well, this Measure 0526, the title is Timely Initiation of Care.

It's CMF. It's a process measure and it's measured at the facility level of analysis. The description is that it's the percentage of home health episodes of care in which the start or resumption date of care was either on the position specified date or within two days of the referral date or discharge date, whichever is later.

So it sounds to me like they would want nothing later than two days, I would guess, the way the measure is described. Now, I have to admit that I had some problems with this measure.

I was confused by some of the date or some of the evidence, I suppose. The importance of the measure, of course, everyone thought it was high. The opportunity for improvement, two of us thought it was medium, a medium opportunity. One thought high; one thought low. I'm not sure how we'll rectify that.

But the thing is, at the end of (1C) where we're looking at the body of evidence, there are three people who thought that, no, it didn't meet the body of evidence criteria and only one felt yes.

Now, if I understand correctly, if we get to the body of evidence and there are -- well, I guess we have to decide whether we want to go on, what the whole group thinks about whether this meets criteria and can move on. But as far as I did, I thought, no, it didn't.

But I did go on and evaluate the risks. Now, do you want me -- I'm not sure; I haven't done this before -- do you want me to go on and talk about the rest of the evaluation or do you want to stop here, talk about this issue right here at the end of 1C, at the end of the evidence (base)?

Karen Johnson: Suzanne, this is Karen Johnson. For the preliminary evaluations, yes, let's just go ahead and finish your summary, yes, and just to clarify the process a little bit, in the in-person meeting if a meeting doesn't pass importance to measure and report, at that point, the committee would pretty much stop discussing that measure.

But for these preliminary evaluations we didn't want you to stop even if you thought it wasn't, you know, shouldn't go on.

Suzanne Heurtin-Roberts: Okay.

Karen Johnson: So yes, so if you're just continue your summary, that would be great.

Suzanne Heurtin-Roberts: Okay, so the body of evidence was in question. We'll talk about that in a minute. As far as scientific acceptable, everyone thought that, yes, it was acceptable, it did meet the criteria. For usability, everyone thought it was high except one person who thought it was moderate and that's certainly sufficient.

And let's see, for feasibility, everyone thought that it was feasible, yes, that it met the feasibility criteria and that was high. So at the end of the judgment of the measure as a whole, three thought, yes, it met all the criteria and one thought no.

Now, I think that one of the critical things that we're going to have to address is the body of evidence. And the issue was that there was really one study that was listed.

Now, let's see. The quantity of the study, obviously everyone thought that the quantity of evidence was low because there was that one study. The quality, three of us thought was medium, one thought was high. I really felt that it was difficult to assess with the information that

we were given, although the sample was, you know, it was an incredibly large sample. I think it was 1.9 million cases.

And then the consistency, I personally felt, was difficult to assess because we only had that one study. Other -- let's see, one person thought that there was insufficient evidence and two thought that there was medium.

I think that since there's only one study it's hard to even think of -- it's hard to think of consistency. So do we want to talk about that at this point? I think this was a big sticking point for most of us.

Karen Johnson: Any other work group members want to comment on that?

Lorna Lynn: I mean, I had to -- this is Lorna Lynn. I thought, like you say -- I agree with everything you said. It was one study but it seemed to have been a well done study with a huge amount of data.

It was -- the results were also interesting. You know, their findings were interesting in that what they found was not what would have been expected and their findings seem to go in different directions.

Suzanne Heurtin-Roberts: Yes, and again, since they went in different directions I really didn't know how to evaluate that. So that's going to be something we'll need to discuss certainly at the larger work group meeting. I don't know how much more we can talk about it here. But I was certainly puzzled by how to approach it, I suppose.

Thomas Howe: Yes, and this is Tom Howe. I think this process measure shares some other characteristics of some of the other measures that, you know, I just want to put it in the common sense category that, you know, if they've ordered a service, it makes sense for that service to be

provided when, you know, experts or at least consensus suggests a reasonable timeframe has elapsed, like two days.

And I'm not sure how much energy you're going to get in the clinical arena to do a number of studies on this. So I kind of gave the body of evidence a pass understanding that the likelihood of there being significant studies is I think is going to be low.

Lorna Lynn: That's a good point. It's hard to think where you're going to get the funding to do that kind of study.

Suzanne Heurtin-Roberts: Yes, and the data source is, you know, that's -- the data source is pretty immediate. It's OASIS, so it's all the actual data of actual cases that's gone -- that goes into this database. So you know, it's not that these were gone on -- that there was -- investigators have gone off and, you know, made a -- done a sample of it's, you know, practical of the entire population, I suppose, of the scientific population of the case in question.

So, you know, the data is seemingly good quality data and there is certainly a lot of it. But, again, it's one study and I hear what you're saying. I'm not sure how to approach that. I don't know how the rest of us think.

(Don Finuchio): Well, I'm not sure that anyone can hear me. This is (Don Finuchio). I'm the person that wrote the study that you're referring to and, indeed, home health in general suffers from a lack of scientific reporting for a variety of reasons, which I won't get into here.

But suffice it to say that virtually -- there is minimal research on virtually all of our work in the area of home health, regardless of which outcome measure from a quality out come perspective and certainly given that process measures are extremely new to our system, only in the last couple of years. There's virtually no research on the impact of process measures.



Suzanne Heurtin-Roberts: Okay.

(Don Finuchio): Regarding your comment, yes, the data are essentially the population of episodes of care that occurred during the year 2001, calendar year 2001.

Suzanne Heurtin-Roberts: Yes.

(Don Finuchio): So we have an excellent representation of that year. What has happened in the past decade, obviously a need to replicate would be appropriate. Regarding the findings, what we found was a decreasing curve that is the number of days was and for the functional measures -- excuse me, for the acute care hospitalization, the two utilization measures, there was a major difference between whether or not this was a start of care or resumption of care.

In general, what we found was that the earlier you got in it made a difference in terms of the functional measures like improvement in ambulation, improvement in bathing. But there was very little difference and almost an inverse difference in terms of the utilization measures for the resumption of care patients.

Resumption of care patients are patients who have gone to the hospital recently and have now come back to home healthcare and are almost, by definition, more fragile.

Suzanne Heurtin-Roberts: Okay, okay, understood. And that actually moves us forward to the next area about scientific acceptability. Pardon me?

Karen Pace: Hello. This is Karen Pace. I'd like to just address your question about the -- to the steering committee in terms of (direct) criteria.

Suzanne Heurtin-Roberts: Okay.

Karen Pace: And you're doing exactly what we asked to evaluate the evidence that does or does not exist but we do an exception regarding the evidence criteria for measures that the steering committee as a group can consider in terms of whether it's something that would be expected by, you know, that there is consensus by expert opinion, for example, and that the benefit greatly outweighs any potential harm.

So I think it's good to -- and you've done exactly what we've asked you to do, to evaluate it based on what was submitted. But I just wanted to remind the steering committee that you will have the opportunity to identify whether you think a measure meets an exception for having that body of evidence.

Suzanne Heurtin-Roberts: Okay. We've heard you, I think. I do want to move on then to scientific acceptability. Everyone agreed that there was high reliability. The reliability studies, you know, definitely showed high reliability and let's see. For validity, everyone thought it was high except for myself, again.

And I put it as medium. Now, when I'm looking at this I think I did that because, again, I was puzzled -- well, I was puzzled by the seemingly contradictory findings of improved function but also increased rehospitalization that was associated with timely transition to home care.

When I'm thinking about it now, I'm not sure that that's a question of validity. I'm not sure where that fits in, actually. But those findings did puzzle me and I didn't know how to think of those. But I would certainly be amenable to moving my score up to high rather than medium.

That being said, everybody decided that there was a high degree or that, yes, it was scientifically acceptable, the measures and the information given. So it passes on that, that criteria.

In terms of usability, these are high or moderate usability. Let's see, only one person -- one person thought that the usability, in terms of public reporting and quality improvement was only moderate. The rest of us thought it was high. But I still think it passes because we had three highs and one moderate, so we all thought it was usable.

Does anybody want to talk about those differences there or the differences -- questions about scientific acceptability?

Lorna Lynn: This is Lorna Lynn. I guess this question is also for our NQF advisors.

Suzanne Heurtin-Roberts: Right.

Lorna Lynn: With the validity evidence that we have, with what we heard about the difficulty in doing research in this population, it seems, certainly based on logic and common sense, that having prompt follow-up home care is a good thing and would do no harm to a patient.

And yet, if we're not -- if we don't have data that there are benefits that you might see longer term in terms of something like, you know, three to six months mortality or decreased hospital days overall, is that something that NQF would still want to endorse in the absence of that kind of evidence?

Karen Pace: This is Karen Pace and I think that's something that we would like the steering committee to provide some guidance on and recommendations. So the question is whether this -- you know, even though there hasn't been studies done, whether this represents something that really should make a difference in patient outcomes.

And I think, you know, we need to -- the question that came up about the validity testing and the correlation, I think it's basically the same study, I believe, right, that was used for both the evidence and validity that ((inaudible)) understanding.

And that's probably just a function of, you know, what they had available because there hadn't been a particular study about this. But I think it's just going to require the steering committee to -- this is an area where your expertise and ((inaudible)) are going to play, you know, a big role in terms of how to view this measure in light of our criteria.

Suzanne Heurtin-Roberts: Right. Well, it's good to know that some of this is -- can be sort of subject to our interpretation of the way we've evaluated things and our interpretation of the information given us because if we do it strictly, for a lot of these measures there would be real problems.

But if there's room for opinion, educated opinion, then I think maybe that's a good thing.

Karen Pace: And we need to follow a specific process. So we first ((inaudible)) evaluate the evidence that does or does not exist and then, you know, decide whether it meets, should be considered for an exception and then the reliability and validity, I mean, the questions about the validity I think are things that need to be raised and, you know, decided whether, you know, the reasons or explanations are acceptable.

So it's really a combination of using both.

Suzanne Heurtin-Roberts: Okay.

Thomas Howe: Well, is it reasonable for the steering committee to have as a paradigm really that where we have a process measure like this that there's a line of sight at least to an outcome result? I

mean, I don't know how we set those parameters or timeframes and we certainly aren't going to fund the studies.

But does the measure steward have some responsibility in giving us at least some idea about how this ends up having an impact on outcomes?

Karen Pace: Yes, and I believe that -- and there's a couple places in the current measure submission that we can take a look at to see what they said and that's under 1B1 where they -- where we ask them to -- what benefits they think this particular measure can have.

And in that evidence section, 1C, we do ask them to start with identifying kind of the structure process outcome linkages, so I don't know -- so basically, for example -- and I don't know, Lauralei, if you want to bring this up on the Webinar so we can see in 1C1.

Lauralei Dorian: Can you see it now? It's up on the screen.

Karen Pace: Okay, right. So they talked about this being related to the outcome, timely initiation of care should lead to quicker identification, resolution of patient problems that drive use of hospital ((inaudible)) urgent care. So, you know, their suggestion is that this would lead to decreased hospitalization in urgent care and I think that's where the question came up about some of the mixed results of the study.

Thomas Howe: Yes, I don't want to beat this up too bad but, I mean, is there an expectation that there will actually be some actual results, not speculation on what the results ought to be?

Karen Pace: Good question. And so what we would expect, for example, when this measure came back for endorsement statements, would be additional analysis on this validity issue on, you know,

does this performance on this measure actually correlate with performance on the outcomes that were suggested that it should be correlated with or some reasonable explanation of that?

So, you know, all the measures that are endorsed do come back through an endorsement maintenance process and, you know, we at that time of endorsement maintenance are expected to meet the same criteria.

But, you know, again, this is something that I think you all need to talk about in terms of whether, you know, this is the right focus for a national performance measure.

Keziah Cook: Hi. This is Keziah Cook. I'm from the measure developer. Would it be helpful if I just spoke briefly to the validity testing that we did for this submission?

Suzanne Heurtin-Roberts: Sure, yes.

Keziah Cook: Okay, so we did -- there were two things we did. One is we looked at the data that, you know, was initially -- was collected at the start of 2010 using the (Oasis C) instrument. And that gave us information about was care initiated in a timely fashion.

And our first piece of analysis was to try to relate that to outcomes and we got some conflicting results, you know, as we displayed on the measure forms.

Then we also convened a technical expert panel -- this was in December of 2010 -- to look at our findings, to talk about, you know, the literature around this measure and to also assess the face validity of this measure.

And the technical expert panel, you know, was presented similar evidence to what we included in our NQF submission, you know, and then held an in-person discussion. And seven of the 10 TEP

members rated this measure as having high face validity and another one rated it as having, you know, partially meeting the validity criteria.

And I mean, we actually had a very interesting discussion with the group and they basically pointed out a -- and this is sort of something hard to get around. With ((inaudible)) we only collect data for those patients who successfully initiate home care. So the assessment data is collected at the first visit.

So if timely initiation of care does not occur, you know, so if it's several days or a week after the patient's maybe been discharged from a hospitalization, the home health agency still hasn't initiated care, if that patient ends up back in the hospital, we'll never see them in our data.

So the technical expert panel basically thought that the lack of relationship we're seeing between timely initiation of care and what everyone agreed was sort of a plausible expectation that that would lower hospitalization was really fueled by the fact that when care isn't initiated in a timely fashion we often aren't collecting the data.

Karen Pace: And I guess the question: do you have any way to quantify to what extent that is happening that patients are, that were referred to home care are readmitted before they get sort of started?

Keziah Cook: You know, I mean, I think from our perspective, you know, as the measure developer using OASIS data, it's something that's, I mean, it's almost impossible, you know. If the care isn't -- you know, it was recommended by not actually initiated, we won't see a Medicare claim for home health services. We won't see an OASIS assessment.

So conceptually one could, you know, do a chart review of a bunch of hospital discharges and see was home healthcare recommended or, you know, prescribed so to speak for patients? Was it initiated and then what happened to those patients?

So, yes, theoretically you could do the study but using our data sources, which are the OASIS assessments and the home health claims, we're not able to do that.

Thomas Howe: Well, I guess I know we have to move on to the other measures but the denominator statement and exclusions doesn't actually refer to folks that should be in the denominator dropping out. So that clarification really needs to get made if that's the case.

Karen Pace: Right because this is our patient population, our patients that have completed OASIS assessments, so, I mean, again that's -- you know, from the measure perspective, our exclusion statement's a complete statement but if you think about, you know, who is this measure capturing, it's not the complete set of patients for whom home healthcare was recommended.

Karen Johnson: This is Karen Johnson. These are really interesting questions and I'd love to keep going with them. If you have more burning questions on this measure, let's go ahead and discuss them now. But if nothing burning then I think we do need to move on to the next measure just because of our time constraint.

Suzanne Heurtin-Roberts: I have nothing else.

Karen Johnson: Okay.

Suzanne Heurtin-Roberts: Everyone agrees that the study was feasible. Certainly the data was already collected, so yes it was. And it's -- most of the group felt that it met the criteria for endorsement. So, I mean, that's the bottom line.

Karen Johnson: Okay.



Suzanne Heurtin-Roberts: Okay?

Karen Johnson: That sounds like we'll be hearing some more interesting discussion in the in-person meeting as well.

Suzanne Heurtin-Roberts: Hopefully we can talk with the others, okay.

Karen Pace: And also just a quick reminder from NQF that you can make use of the SharePoint discussion forum ((inaudible)) discussion.

Suzanne Heurtin-Roberts: Okay.

Karen Johnson: Okay, then the next measure on our agenda is 0511 and we had asked (Mark) to take that measure as lead discussant. (Mark), are you on the line? If you happen to be on the line and we just can't hear you, can you hit star 1 so that the operator would know you're there? Okay. (Mark), are you there?

Operator: This is the operator. I'm not seeing anyone queued up in the queue here.

Karen Johnson: Okay, thank you. Would anybody else...

Female: You could go on to the next one and then maybe come back to that unless there is another reviewer that wants to take it.

Karen Johnson: Yes, I was going to let somebody else give it a shot if they want to or I could give it a shot.

Female: Okay, go ahead.

Karen Johnson: Yes, let's just go ahead and do this one. This one is correlation with an existing imaging study for all patients undergoing bone -- I don't know that word -- scans.

The measure is percentage of final reports for all patients regardless of age undergoing a bone scan that includes physician documentation of correlation with existing relevant imaging studies such as x-rays, MRIs, CTs that were performed.

So just going through the results here, importance to measure and report, there was a pass with one yes and two no. And for impact, two people thought it had high impact. One person thought it was medium impact.

Performance gaps, one high, two medium and then for evidence it was one yes and two no on passing that one. And the problem there was the quantity of evidence and, of course, consistency as well since there was little quantity. So it's kind of a similar problem as what we've already talked about, there's not -- the body of evidence, again, is one of the comments that was made.

For scientific acceptability, that one, there was three yes and no nos with reliability getting three votes for medium and validity getting two medium and one low.

And the main comment on that one, I believe, is the concern that the physicians could use exclusions from, I believe, the denominator is reasonable effort to locate other studies weren't successful.

But what constituted a reasonable effort was not well-defined, so that was one of the main concerns and just the whole idea that that is a very subjective kind of way to measure that.

So usability, one high, two mediums; feasibility, same one high, two mediums. So in conclusion, two -- whether it was passed or not or being recommended or not, two yes and one no and, again, I think that the main comment was whether there was reasonable efforts and what that might mean and would that be consistently applied.

So with that really short summary of the preliminary evaluations, I'll open it up to see if anyone wants to comment on any of those concerns.

Lorna Lynn: So this is Lorna. I think you did a great job summarizing both the measure and the comments that were there. This issue of reasonable efforts was, for me, a key thing and I wonder if the measure developers, measure stewards, can help us with that and if this is something where we can go back to them before we meet together to see, you know, is there something more that just we're missing from the application that could help us understand what the expectation is?

Joanne Cuny: Yes, this is Joanne Cuny from PCPI; just confirm you can hear me.

Lorna Lynn: Yes.

Joanne Cuny: Okay, great. Diedra and (Mark) may be there as well. But in terms of what we found on the testing project that we did, there -- as you might see, there were of the 97 records that they looked at, there were 18 that had documented exclusions or exceptions and so those patients or charts, records were removed from the denominator.

But we have a verbatim report so that we could actually drill down into what was actually happening with those 18 and so that there is documentation of why they were removed.

And in -- so in two of them, it's very difficult to see anything other than there was nothing available or prior exam, no longer available, so you really can't tell exactly what happened in there.

But in 16 of the 18, they say that there were no comparative x-rays or there were no appropriate studies for comparison and we did hear from the professional data extractors we used at (QIO) for this project that in some cases the bone (sintigraphy) that was ordered was on a metastatic process so that other exams that would be available for that patient would perhaps be of the lunch instead of the bony prominence that was being studied.

So that's -- I guess that's the more that I would tell you that we see in the reports here. So in 16 of the 18 that's the type of documentation here that the prior studies were not relevant to this new study.

Thomas Howe: All right, I shared the same concern that the reasonable availability and subjectiveness disturbed me as a denominator exclusion. I do get it about relevant. I'm having trouble kind of getting back to this denominator why it got exclusions in it at all.

I mean, you would reset, you know, the benchmark, but if we're really after coordination of care, it seems to me that the numerator and denominator events are actually pretty clear.

Diedra Joseph: This is Diedra. I am at PCPI. May I make a brief comment?

Karen Johnson: Go ahead.

Diedra Joseph: So I know that in the original review of the measure the same issue was brought up and we did take it back to the work group for further discussion and then intent of the measure, the work group decided that the intent of the measure is to encourage correlation with the existing imaging study.

However, the expert clinicians on the work group noted that or confirmed that there were frequent instances in which existing studies are not available and so they landed on keeping the denominator exclusion or exception just to make sure that we didn't penalize clinicians for not being able to obtain a previous study if they made efforts.

Karen Johnson: Does anybody on the work group want to continue -- response or anything you want to add to that?

Sue Abreu: Well, I'm not on the work group. I was on the original work group for this measure if that's -- I would provide some input.

Karen Johnson: Can you tell us your name just so that we know who you are?

Sue Abreu: This is Sue Abreu. I'm a nuclear medicine physician. The challenge with this is that, you know, we're a referral specialty, so we don't control a lot of what comes to us.

We make an effort to get the outside films but a patient may be seen at institution X for one thing and then come to us for institution Y and not everybody has electronic imaging and sometimes patients check out their films and then take them to doctor in institution Z and you can't get them.

And so we needed a way to deal with that scenario. You don't -- as was stated, you don't always have imaging done before hand. This might be the primary imaging modality and then there are secondary images taken to correlate once you've identified a sight of abnormality.

So it's, you know, sometimes there's a cart and a horse problem, too. And so I would entertain suggestions on what type of phrasing -- it would be very helpful for us to get some feedback on what type of phrasing on reasonable effort might be suggested.

Lorna Lynn: So this is Lorna Lynn and I certainly understand that, you know, the issue of institutions X, Y and Z. However, if we're trying to really address coordination of care and move the entire field forward, maybe it just isn't acceptable. You know, maybe this is something that can push institutions towards a different kind of accessibility of those records and no one is demanding 100% performance.

It certainly is one thing if there are no studies to correlate. That seems like a very different thing than we know there are studies out there but we can't access them.

It's not blaming the nuclear medicine physician but saying something about what he or she was able to do given the situation.

Sue Abreu: I think if it were clear that there would be no punitive impact from any of these measures people would be more comfortable with that. But to be honest, out in the imaging community, that's not the perception of what the long-term issue will be.

You know, there is concern that these would be used in a punitive manner in something that you can't control. But I agree. If it's made clear that we say, look, we want to identify issues and we want to have pure data, you know, and so the only exclusion would be no prior studies done, no relevant prior studies, you know, done at all, that would make sense. That would be a great way to do the measure with less ambiguity and it would help push the world in general, which would be helpful.

Thomas Howe: I would recommend that. I think we do get into the issue of who's measuring whom and to the extent that nuclear medicine doctors feel like this is pointed exclusively at them, I can understand the cautions around the exclusions. But I think that utility from a societal point of view is that this is a measure of care coordination and not really solely aimed at nuclear medicine.

Sue Abreu: I think that would be a -- I don't think there would be a problem revising it to do that as long as there was a very clear statement from the reimbursement world that it would not be used to penalize -- it could not be used to penalize.

Thomas Howe: Well, I don't think you're going to get NQF to be able to...

Sue Abreu: No, but that's the problem. I think that's a real world problem that has to be addressed when you want to make these measures -- that's the essence of a lot of the problems. We want measures that are good data but at the same time they're being used for other purposes.

Suzanne Heurtin-Roberts: In describing this measure and submitting information about this measure, can't it be phrased as we're looking at care coordination in terms of systems, coordinating systems? It's not that we're coordinating or we're looking at care coordination in terms of particular professions or a particular individual, a practitioner let's say.

But we're actually looking at the linking of systems. If that's said, can that help the cause at all?

Sue Abreu: Again, I think the critical part is how the data is used by people like CMF. That's the problem. That's the real world problem that we have with buy-in with our providers is a fear that it would be used in a punitive way on something you cannot control.

Karen Pace: So this is Karen Pace from NQF and so I think the committee has identified the question that you're going to have to grapple with in terms of the validity of the measure score in being able to really make accurate conclusions about quality.

So I think the question you've raised is if you have two exact scores do you know that it's because they're doing equally well on obtaining and correlating those reports or the prior study or is it because they have more exceptions?

So I think, you know, it's going to come down to the steering, you know, in terms of, you know, the steering committee having more discussion about this and the impact on what conclusions you can make from the performance score.

Sue Abreu: Yes, I think the critical thing is what do we want to measure: the systems, the care coordination across systems or is this an evaluation of nuclear medicine providers?

Karen Johnson: Okay, well, that was some really interesting discussion there. So I think with that, if everybody's agreeable, we'll go ahead to the next measure, 0645, biopsy follow up. And Bonnie, I believe you were the lead discussant for that measure.

Bonnie Wakefield: Yes, thank you. So this measure is biopsy follow up submitted by the American Academy of Dermatology. And the measure is the percent of patients who have had a biopsy whose results have been reviewed by that physician and then communicated to the patient's primary care or referring physician and to the patient.

They also include documentation of that communication in a log and the patient's records. So it's a process measure. The only exclusions to the denominator are patients who don't have a biopsy and the level of analysis that the clinicians group practice and individual levels.

So there seem to be some disagreements on this measure in terms of the importance to measure and report. In impact there were three reviewers scored that as high and performance gaps two reported that as high and one as moderate.



They don't -- and I think, you know, going back to this issue of evidence, there isn't a lot of evidence for this measure. As someone noted early, this might fall into that common sense category that measure, you know, if the biopsy is done the results, you know, should be communicated to the referring physician and patient. So and this would affect a large population of patients.

So that result was in sort of quantity of evidence, two moderate, two lows; quality, two moderates and two insufficients; and consistency, two moderates and two insufficients.

So there is disagreement on the quantity, quality and consistency of the evidence because there isn't much. So in terms of importance to measure and report, there were two yes's and one no. Fourth person didn't vote on that.

In terms of scientific acceptability of the properties, there was more disagreement, reliability, one high, one moderate, one insufficient and then validity, one high, one moderate, one insufficient.

And I think part of that disagreement is, again, you know, for this kind of measure which seems rather intuitive on the surface but it doesn't have a lot of testing. They've done some testing but acknowledge that it's in a small sample.

In terms of usability, again, there was some disagreement, two high, one low, one inconclusive. And I think one of the important points one of the reviewers raised here was there was no time element, which I'm assuming means, you know, there was no timeframe within which the biopsying physicians, you know, need to report that back to primary care or referring physician, and so that might sort of beef it up a bit.

Again, feasibility was all over the place, one high, one moderate, one low and one -- I'm blanking on what I was -- insufficient, sorry. I think a couple of things here, again, the small sample size.

And this came up earlier sort of requiring this log of contacts and I should have mentioned that earlier that using a paper log, if you have an electronic record, may not be, you know, the best way for the physician to document those results were communicated.

But it's unclear that all physicians have access to an electronic records document. So I think in that case a paper log is the only choice.

And then finally, preliminary -- the preliminary assessment, again, was split two and two and one of the comments there I think that might have sort of lead to that is so assuring follow-up care that's determined appropriate after a test, you know, is of more importance than simple transmission of the results, which is true.

But, you know, it's necessary -- it's sort of the necessary but not sufficient. It is necessary to communicate the result first and it's intuitively sort of makes sense. So I'm not sure, you know, if this person was sort of looking at the follow-up as maybe that should be the measure or if that was implied in the measure.

So I think that covers it.

Karen Johnson: Okay, thank you, Bonnie. Does any one of the steering committee members want to add anything to that? And go ahead and get the discussion rolling.

Thomas Howe: Yes, hi. This is Tom Howe again. I made similar remarks about the first measure in that -- and in this one, I mean, if you turn if on your head, on its head, who would want to do a study that looked at individuals that didn't get biopsy results returned?

I mean, that's just nonsensical. So the evidence base is going to be send because no one would want to set up a study that would look at sort of the negative population that didn't get their biopsies back or set it up so that they didn't.

Bonnie Wakefield: Right, good observational study. But that would just validate that it's occurring or not occurring, but, right, you wouldn't want to do a clinical trial on it.

Thomas Howe: I think some of the discrepancies here may tie back to since this is such a common sense process event is this -- I mean, what are we measuring it for I guess is the question? Is there an opportunity here that really should be out there? And I think that's some of what's coming across.

Bonnie Wakefield: Are you suggesting we shouldn't be measuring it? Is that what I...

Thomas Howe: Well, I see that there is some -- actually, I'm one of the ones that thought we should continue to support this. But I can see that there's quite a different set of -- there's quite a diversity here.

Bonnie Wakefield: Right.

Lorna Lynn: This is Lorna Lynn. I was concerned about the lack of timing. I think it was specified that there needed to be a report back to the patient and the referring physician in the measurement year and that just seemed to lack the specificity that would be appropriate.

And I understand that given these are looking at biopsies there can't be a -- it would be very difficult to have a fixed period but the measurement here seems to be somewhat problematic.

I also was wondering why this was specified to the biopsying physician who needed to give the report back to the patient rather than referring physician. I think one could argue that the physician who's providing the long-term care for chronic condition, for example, would be more appropriate in terms of patient centeredness in transmitting the information to the patient.

Bonnie Wakefield: Which is some -- you know, many tests get communicated back through the patient from the referring physician, correct.

Suzanne Heurtin-Roberts: This is Suzanne. For a number of points I thought that there was insufficient evidence but I think, again, this goes back to the first measure, some of the discussion of the first measure on timely beginning of home health care.

There's ((inaudible)) it seems that, you know, it's intuitive that this should be, should be, this should make sense. So there's not a whole lot of information there.

The question is then as reviewers how strictly we're supposed to follow the instructions to, you know, to evaluate this on the evidence given because I think that the evidence given, even though there's reasons that there's not a whole -- very strong body of evidence, even though I understand, Tom, that there are reasons that there wouldn't be more information about this, you know if we just evaluate this on the information given to us, to me it seems that the information is insufficient.

So I guess the question is do we decide that at this point or is that -- I guess that is something for discussion with the larger committee later. But I had that problem about how strictly to follow the guidelines.

Karen Johnson: Right. I think it's like Karen Pace mentioned earlier. We would like you to follow the guidelines just so that everything is very transparent. But then the guidance also allows for that exception, as she discussed.

And what I'll do in the next couple of days is to send out some -- a reminder about that particular point specifically because I think it could come up on several of these measures and maybe that will -- maybe when you see it again written down it might make some more sense to you.

Suzanne Heurtin-Roberts: Okay.

Karen Johnson: Do we have anybody from AAD who wants to respond to any of these concerns?

Dr. Oliver Wisco: Yes, this is Dr. Oliver Wisco. I'm one of the measure developers. I do really appreciate all of the comments and in developing the measures we actually had the very same concerns that the reviewers had.

There is -- I know I'm beating a dead horse here. There's very little data out there or few gap analyses that have looked at whether our processes for a care coordination are doing -- being done correctly.

This measure was developed through the notion that a lot of tests are redone or duplicated. We see this in the hospital setting for a patient transitioning over to the outpatient setting or simple tests such as the CBC, (Chem7)s, LSG and they are redone because there's not appropriate coordination with hospital physical to the primary care physician.

When there was a call for measure development on care coordination and looking at what dermatology could add to the measure field, we looked at what we did the most for looking at lab tests and things that we need to coordinate with primary care physicians.

And essentially any time that a physicians, not just dermatology, cut something out, we proposed that there is a deficit between the communication between the biopsying or the surgical physician and primary care physician much like that in the instance of the hospital and the primary care physician and the core coordination of test results.

So this measure was developed with the understanding that there was limited data. But also, when I proposed this measure about three or four years ago to the NQF, one of the important things that the role of NQF now, what we felt, is by allowing some of these measures with little data, the development of the measure actually allows for the development of the data understanding it's nice to go and have the data before hand but because everybody is so new at this process the data is just not out there.

And it was hoped that -- or it was intended that measures like these would start opening the door for gap analyses of process, for process improvement. So that's where we come from.

And I agree with you completely. There is no data stating that whether -- that physicians don't get information to primary care physicians in regards to biopsy results. There is absolutely no data.

However, this measure will start helping us develop that data and also start then with the intent of help improve the processes. So I agree, we're going backwards on here but we're also limited in that there's very few things that we could propose in this setting.

To say that a time endorsement is very important, I agree with that completely. But and whether or not the data actually or measuring at the point of communication of the physician and to patient care, we felt that this was -- that step would actually be the step to measure to determine whether an appropriate process was being made.

We felt that this was more the gatekeeper of the process and that's why we chose that part for communication for the point of measurement.

Alison Shippy: Hi. This is Alison. I'm from the staff out at the Academy and I just wanted to kind of piggyback on what Dr. Wisco said. You know, the lack of data and, you know, I think we recognize that to be a parent in the application that we had a difficult time kind of getting larger numbers for our testing.

And I think this was just an unfortunate kind of timing issue. You know, we have included this, or CMF, rather, has included this in the (PQRS) program. So, you know, I think that we are eager to see what kind of results we'll get from a more wide implementation of this measure, so I just wanted to add that piece, so that is in for 2012, the (PQRS) program.

Karen Johnson: And this is Karen Johnson from NQF and Karen Pace can correct me on this if I am saying it wrong, but I believe the time limit of the endorsed measure, what we wanted the developer to do was submit reliability of ((inaudible)) testing results.

But at this point, we don't necessarily expect them to be able to present performance gaps results. So am I saying that right, Karen?

Karen Pace: Well, actually, we don't expect that they would have data from the measure as implemented since it hasn't been implemented yet. But we would expect for any measure some discussion of performance gap, hopefully some data from prior studies or for pilot studies.

It looked like that section of the form was left blank. And as you saw, they talked about attempting to do some measure testing but I don't see any results. So is the measure developer -- you just said inconclusive but you didn't actually give us any of the results. Do you have results from those three sites where you actually did some testing?

Alison Shippy: I think I -- this is Alison again. I think that I had sent just (inter rater) reliability to Karen after the fact. I think that we had discussed that on the staff level that we weren't really sure kind of how to proceed or how to interpret kind of the work that we had done.

Dr. Oliver Wisco: There was some -- this is Dr. Wisco. There was some confusion of specifically what was needed. So just to reiterate what you -- or just to rephrase what you had asked, are you asking for data showing actual performance, so meaning assisting physicians reported. What percentage did they actually qualify for the measure -- actually not -- yes, I guess meaning they fulfilled the measure requirements? Is that what you're asking for?

Karen Pace: Well, that's what we're asking for under performance gaps for measures that were previously endorsed. What's the actual performance on the measure as specified?

If it's a new measure, often the data on performance gap is from prior studies, not necessarily on the new measure. In reliability and validity testing, we're actually -- we're asking for the data. So if you did -- if this was medical record extraction, what was the (inter rater) reliability on your reliability testing? And I'm not sure what you were doing for validity testing.

Dr. Oliver Wisco: We do have the (inter rater) reliability testing and, in terms of how we actually performed, we do have that data as well and we can get that to you. And once the -- and we apologize if that information was not clear. We were a little confused in what we came on for getting this data together. Yes, we did not have everything ready to go just because of transition of staff.

So we can get that to you. In terms of validity data, there was, once again, there was some issues there and we can get that as well.



Karen Johnson: So if you could get us that as soon as you can, we will definitely get that out to the entire steering committee so that they can evaluate this measure a little bit more fully.

Karen Pace: Right and one of the things that we'll do after this call is it'll probably be best that we actually have them enter that into the form where it's supposed to be so that everyone can follow along more easily. So we'll follow up with them.

Karen Johnson: Okay, we'll definitely open up that platform for you so that you can do that. We're really almost out of time for this measure but were there any other questions that you wanted to pose to the developer before we go onto the next measure?

Okay, if not, let's go ahead now and skip measure 0494 for the time being and go on to measure 0553 and, Lorna, I think you were to be the discussant for that measure.

Lorna Lynn: Right, so this is a measure that comes to us from NCQA, which is a medication review in older adults. The brief description of the measure: it's a percentage of adults 66 and older who had a medication review including prescription, over the counter and herbal or supplemental medication by ((inaudible)) practitioner or clinical pharmacist during the measurement year.

In terms of the important to measure and report, the impact was rated high by all four reviewers. Performance gap was rated as high by two and moderate by two. NCQA did provide us with three or four years of data that interestingly showed improving performance over time, which is a nice thing to see.

With regards to the evidence for this measure, we were kind of all over the board on this. One rated the quantity high, two medium, one low. Quality was rated high by one, medium by two and low by two. Consistency rated high by one and moderate by two and low by one.

There were quite a few studies that were cited. Older studies appeared to one reviewer to be stronger but there was -- I noted that this measure remained an important emphasis by both the Institute of Medicine and the IHI.

Systematic reviews are acknowledged by the measure stewards who have shown mixed results for the effectiveness of the process. One other reviewer did not that it's hard to see any potential harm to patients in terms of a review of their medication while it's easy with common sense to perceive significant potential benefits.

With regard to the reliability and validity of the measure, the reliability was rated high by one, moderate by two, insufficient by one. Validity rated as moderate by three. I think we were missing some information about the validity testing that perhaps someone on the call can give us some more information on.

It was described as being limited to face validity testing with experts but we didn't here -- there was not a whole lot about that in the application.

In terms of usability and feasibility, these were passed by all the reviewers, although mixed ratings of high and moderate on these. And in terms of a preliminary assessment of its suitability for re-endorsement this was all four reviewers said yes.

Karen Johnson: Okay, thank you. Does anyone on the steering committee have any other comments they wish to make before opening it up to overall discussion? Okay, do you guys want to discuss the concerns that came up that Lorna summarized for us?

Female: Hi, this is the NCQA team. Can you hear us?

Lorna Lynn: We can hear you, yes.

Female: Okay, great. So trying to -- you asked for some more discussion about the validity testing and reliability testing.

Lorna Lynn: That would be helpful, particularly in light of the comment about the systematic review yielding different results.

Female: So, you know, this is -- we've had this issue with, you know, several of these measures in care coordination in terms of there not being evidence in the same way that there may be for, you know, guidelines because people don't do necessarily (RGT)s on medication review. You don't want to, you know, randomize someone to not get a review.

So in terms of the systematic review and the evidence being mixed, a lot of that has to do with the quality of the studies that are out there. You know, it's been widely acknowledged that medication review is important. The (ION) has repeatedly cited that as an important process that can reduce adverse medical events.

However, in terms of studies that have looked exclusively at medication review, and that's medication review not combined with other sorts of care coordination activities -- and I'm speaking here as like (Eric Coleman)'s work, etc cetera, where medication review is one part of care coordination intervention.

So there haven't been that many studies that have looked exclusive at medication review. And then beyond that, medication review is something that varies in terms of what do you define as medication review in different studies.

And so when we cited that the instances mixed, most of that had to do with the fact that the evidence is not really -- there's not a robust body of evidence. However, that does not diminish

the importance of this and the general expert opinion that this is a -- definitely the benefits outweigh the harm of doing it.

It's an important step, particularly for older adults who are far more likely to have adverse medical events. So does that -- in terms of the testing results -- so for the -- sorry, one moment. I'm just looking over something.

So in terms of validity, we've done a lot of testing on this measure. It's been in our measure set for quite a while. We have done lots of testing in terms of comparing multiple data sources to make sure that this measure is accurately getting at medication review.

And we've also done testing of reliability using a data binomial model which looks at a signal to noise ratio. In general, when we talk about the validity in the (NQS) forum, we talk about face validity because we have such a rigorous process here for measure, creation and maintenance. So this goes through many, many committees of experts, both health plans, providers, measurement experts and then it's open to public comment.

And so, you know, when you say, oh, it's just face validity, that's not necessarily -- I mean, we have such a rigorous process here that it really has been put through the ringer.

Bonnie Wakefield: I have a question for ((inaudible)) yes, because I did see that you, you know, you said that there was face validity. But my concern wasn't that you said there was face validity. But I didn't see -- I don't have the measurements in front of me. I can't tell right now. But I don't think I saw any description of how you assessed face validity.

Female: So I'm sorry. Face validity was assessed through a work group expert panel and the additional information in the measure, AD1, we have a list of the names of people -- the members of our panel, which is a geriatric specific measurement advisory panel.

So the measure was reviewed for face validity by that panel. They all deemed the measure to be valid and important. And it then went on to an additional committee which all our measures go through but also of health plans and stakeholders, physicians that also deemed the measure to be valid and important.

Karen Johnson: This is Karen Johnson. Just to reiterate some of the NQF items on this, we understand that you did face validity and it sounds like you did it in a very systematic way.

I think that the two things in that section that we were looking for really was -- or really the main thing that we were looking for is the testing results, so not just the statement that you found it valid but, you know, did X percent of your committee find it, you know, an excellent measure, you know, that sort of thing?

That's kind of what we were looking for in that and I think that might have been why some people felt that maybe this -- there was insufficient information in your submission.

Female: So I am happy to provide that information in terms of the voting results. All of these measures are voted on and needs a majority to move forward. However, almost all of our measures there is complete consensus among the committee members and I can get you the specific information in terms of how the vote on this measure broke down.

And we apologize. That information has actually never I think been requested before in an NQF forum, so we are happy to provide that to you.

Karen Johnson: That would be very much appreciated. One other thing that -- just that I had a question about and actually included I think in the staff notes was your reliability testing. And you had

mentioned that you had done a signal to noise analysis and the reliability was calculated at this point, 98712. We were a little unclear about that. Is that the average reliability?

Female: So, you know, the data binomial model -- and I don't want to go into a whole statistics lecture here on that model -- it's the model that we use here because it's particularly well-suited to a measure which is binomial, meaning yes and no and there's no distribution.

What you can -- that number can be compared. It's very similar to something like an inner-class correlation (co-efficient), so it ranges from zero to one with one being the highest. So that reliability co-efficient looks across all of our plans and looks at within plan variation and between plan variation, so that's where that number comes from.

I'm happy to provide more details on that number if you would like. I can prepare a memo on how that number is derived.

Karen Pace: This is Karen Pace. (Those) can follow up with you but generally on that method there is a reliability statistics for each measured entity, like the physician level. And so we were just wondering what the distribution was or if this one number that you gave was the average out of everyone.

Female: This number, we only specified this measure at the plan level.

Karen Pace: Okay, and how many plans were actually in the analysis then?

Female: Let's see. This is s a Medicare plan. This would have been probably -- I believe we have all Medicare advantage plans or subsets. I believe it's around 270. I can -- or it might be actually -- I can find that exact number for you if you'd like.

Karen Pace: We can get that for you but you just now mentioned this is a plan-level measure but your level of analysis you have it down to clinicians level.

Female: So this is a continuing, you know, something that we continue to work with NQF in terms of what's the right way -- what are the right boxes to check. We specified this measure at the plan level. We do all of the testing and validation at the plan level.

However, plans often take our measures -- and this applies to all ((inaudible)) measures -- and sometimes filter them down to the individual physician or practice level. So plans are using this measure on a practice or physician level. However, that's not how we specify the measure.

So we don't have any data on that. So, you know, and this is something we're happy if NQF would, you know, prefer to mark this as just a plan level even though it's used on multiple levels. That's fine with us.

Karen Pace: Well, that's an area that NQF is looking into some clarification. So thanks for that.

Female: Okay.

Lorna Lynn: I have one more question for NCQA. This is Lorna Lynn. You made a note that you allowed health plans for optional exclusions to their results. Can you explain what that would mean with regards to this measure?

Female: Hold on one second. I'm going through it. Oh, by the way, I just found it in the report. The number of health plans is 316, just to go back to that. Okay, so which are you -- which question are you referring to on the report?

Lorna Lynn: This is in 2B3.1: NCQA currently allows health plans for optional exclusions to their results.

Male: Hang in there.

Female: Hang in there, sorry. So I'm going to let one of our other people in the room answer for me on that.

Male: ((inaudible)). I don't see anything. This may have been -- this may be, for this particular measure, checking the wrong box. We're just verifying that. We thought that -- many of our measures have optional exclusions. But in this case, I don't believe that that's correct, so that may be a correction we need to make.

Female: That may be a correction we need to make in the report. There are no optional exclusions for this measure.

Lorna Lynn: Good to hear.

Karen Johnson: Are there any other concerns or questions that you want to ask NCQA? Okay, if not, that actually does leave us a little bit of time to talk about measure 0494. That measure, as you now know, we had lots of difficulty with that measure for a variety of reasons, including some technical difficulties at the last minute.

So we weren't able to get out the submission to you in time to do evaluation. But hopefully you were able to take a peak at it and maybe would have some questions for NCQA.

And just to kind of start us off, I might ask NCQA reps maybe just take, if you can, maybe two minutes, just very, very briefly to describe what this measure is and the fact that it's a ((inaudible)) measure and that sort of thing and then we'll open it up for any questions that the steering committee members may have had.



Female: Thank you for that opportunity. If you read over this form, you'll notice it's a very different type of measure than the other measures that you're seeing here. This is a -- the medical home system survey is what we use here at NCQA to -- for recognition of the patient-centered medical home.

And we are putting this for re-endorsement, continuing to make this publicly available. We use this form to recognize health practices that are implementing practices, both process and structure, consistent with the patient center's medical home model.

We have over 6000 practices and physicians currently endorsed by us. And we -- this model is built on ((inaudible)) on the chronic care model, which has quite a bit of evidence. We are also beginning to see evidence specifically of the NCQA-recognized model.

There have been some studies just starting to come out about the effectiveness in terms of improving care outcomes and reducing costs.

One thing you may notice is that the testing data is very different here because this is a non-self report measure. This is a measure that required documentation that is reviewed here at NCQA to make sure that the physician report is backed up with appropriate documentation.

So I'm going to leave it there and then open it up to questions.

Suzanne Heurtin-Roberts: Well, I guess when I did take a peak at this earlier -- this is Suzanne -- and I was puzzled because it looked to me rather than being a measure it looked like it was a lot of measures and rather that we're looking at a survey which seemed to me to be a tool.

So just from, I guess, the developer's end and also from NQF, why was this submitted as a whole instead of each measure being evaluated?

Female: So this is NCQA. We have had extensive discussions with NQF going back and forth on this measure. This measure was previously submitted as a whole so we thought with the re-endorsement we were only asked by NQF to submit this measure as a whole.

Also, these pieces are not meant to be used individually. They really do reflect a holistic concept. And so, you know, I think that one way to think about this is that it's kind of similar in some ways to some of the (cap) surveys where you have multiple items. It's covering a wide range of topics. However, you come up with, as in the case of this one, six scores.

Suzanne Heurtin-Roberts: Okay.

Karen Pace: So this is Karen Pace. So I think you raised some good questions. This is kind of a -- something that doesn't fit into our normal forms and so we are -- have been trying to work with NCQA on this.

And the first clarification is that NQF does not endorse the survey. So we did ask that -- and I haven't had a chance to look at this recent ((inaudible)). We asked them to be very specific that they were putting forward for consideration discrete measures that should have very specific specifications in terms of how those measures are computed, what data goes into them and it still has the same requirements for evidence.

It's just now they have to look at the evidence for all of those in one form versus five or six forms. So we'll be going through this in more detail but as, you know -- and we understand that it's different than anything that you've looked at so far, but just to point out those couple things and hopefully we can work through this to identify ((inaudible)).

Female: And I will clarify that this is a measure that is computed through a survey, similar to the way that the (cap) measure is computed through a survey. So you'll see lots of details about a survey and it is called a survey, however, we presented data on six measures, which each individual measure has the testing data in the form as well as evidence.

Lorna Lynn: This is Lorna Lynn, just a couple quick clarifications. It's a survey but NCQA requires documentation to support the answer to the survey. Is that correct?

Female: Yes.

Lorna Lynn: In terms of screen shots or practice processes to deliver, you know, such as in the improving performance type of measure?

Female: Yes, there is extensive, an extensive list of the types of documentations that we ask for and we worked with practices to come up with the right documentation. So that's all included in the attachment labeled Specifications.

Lorna Lynn: I know that when practices apply for the medical home recognition it's a good bit of work and often can be transformational for the practices as they go through what they need to do in order to gain recognition.

Do you have any data? I would imagine the answer to this is going to be no. How would you have data from any practices who haven't submitted their evidence to you? So there's a little bit of circular logic in this.

Female: You know, we can't -- we don't have data on the practices that don't submit data to us. We can only -- you know, and what we've shown in here is across and in the performance gap section

you'll see that performance on each one of these measures has improved over time quite significantly.

So the plans are getting -- there's been more implementation of the individual items within this and general improved performance and our plan is growing substantially, so, you know, we have over 6000 recognized and we are anticipating, I believe -- what is the number that we're in? It's something like 2000 more this year, so this is definitely something that is (worthy).

Emilio Carrillo: Hi, this is Emilio Carrillo from New York Presbyterian and we have done 14 practices and received the accreditation, the designation rather, and it is a long process. It's very specific. There's requests for a lot of information and then there's a lot of back and forth for weeks and weeks to clarify the backup information and to get, you know, get really to a very, very high level specificity. So I can vouch for that and speak to it.

Female: I'll also state that we have several practices that come back for re-evaluation to improve their scores, so we are seeing improvement in that realm as well.

Emilio Carrillo: And every three years you have to resubmit.

Thomas Howe: Yes, I have a question about -- I guess it's an NQF question. I recognize the measure steward is NCQA and the survey is from NCQA but I think we're looking for a generalizable measure that other entities using some other process to attain this -- the assurance that these six elements are being met by a practice would qualify, that they would be in the numerator.

Is that clear or does this measure hinge on that NCQA survey?

Karen Pace: Hi. This is Karen Pace and this is just another clarification. Unlike (cap), this survey is really a data collection tool, standardized data collection tool because it's not asking patients for

information. It's asking for actual data to be submitted by the practice or combination of things of that nature.

So it's a little bit different than a typical survey where you -- but nonetheless, it is a standardized data collection tool that has been put together by NCQA, so you're right about that.

In some regards, it would be similar to -- we have some measures that are based on a registry data collection where they have a standardized data collection system to submit data for those measures or potentially some of the CMF measures using an OASIS data set or the nursing home (MBS) data set where they have very specific definitions and data collection systems.

So, you know, that doesn't preclude NQF endorsing a measure but you're absolutely right. The way the measure is specified is tied to that survey and I think it's, you know, a good question for your discussion among the community of, you know, whether this could be used outside of that context or, you know, and have some discussion about that.

But the fact that it's -- their survey doesn't preclude that from your consideration.

(Bob Rim): This is NCQA. This is (Bob Rim). And I ask my other team members here to add to this. It's correct that while we do have a recognition -- a formal recognition program that involves NCQA's review and basically interactive work with the practices and clinicians who are seeking recognition, we are making this available publicly through NQF endorsement to make it available to anybody who would like to use this tool and this survey process.

And I think maybe give you an example, a plan may be in their mind's eye thinking, gee, I'm seeing a lot of people go through the NCQA survey on their own. Maybe I want to do that; maybe I don't. Gee, it's available here, NQF. I can look at the specifications. I can use that as a template and I can basically adopt this as a measure and see how my practice is doing.

And it doesn't require any interaction with NCQA, not proprietary and I think that that's the generalizability that we're trying to get to.

Yes, there is a -- you can seek a formal recognition through NCQA. Or if you want, you can use the, you know, publicly available survey tools that's available on the NQF Web site and, you know, and do with that as you will as long as you follow, you know, the agreement that we have in place as measure steward.

You can't call it an NCQA recognized practice but you can be, to yourself, a self-evident practice that does better than your practice down the street. You know, that's up to folks how they decide they want to place (title) on that.

But in truth, anyone can play in a level playing field. And so far, it's access to the survey.

Karen Johnson: Okay, thank you for that discussion. That was some really interesting questions and it was good to hear NCQA's viewpoints on that. With that, I think we're going to go ahead and stop discussion on the measures now and open up the line for the public. So, operator, can you open up our public line for comments?

Operator: Sure, please standby. Okay, all lines are open at this time.

Karen Johnson: Okay, do we have any comments from the public? Okay, thank you, operator. If not, I think nobody may object to ending the call a little early. Before we do that, we do want to give you some upcoming steps so that you know what to expect from us. So Lauralei, take it over.

Lauralei Dorian: Yes, thanks, Karen. Thanks, everybody, for your participation today and thanks to our lead discussants and the developers for being on the call, especially those who have been on two two-hour calls today. We appreciate that.

So we at NQF will work on summarizing these two work group calls, which we'll send out to you. We do want to remind you that there is one more work group call tomorrow and you're welcome to dial into that call either to participate or just to listen to the discussion.

You're also welcome to go back to the SharePoint site and continue any conversation you have on any of these measures and you can also go back and vote on the measure in the work group tomorrow or revote on these measures, if any of the discussion today has sort of ((inaudible)) about that.

So we also wanted to mention that our in-person meeting is a two-day meeting and, as you know, we haven't received any new measures, so we'll probably have a bit of extra time.

So we've been talking with our co-chairs, who unfortunately couldn't be on this call today. But we were talking about how we could best utilize that time on the second day. And if you recall, during our in-person meeting in October, we talked about the 25 preferred practices that were endorsed in 2010 in Nicole McElveen's project.

And we were thinking that it would be a really good exercise to have you review those measures and think about the best way that they could sort of be used moving forward to capitalize on all of that work that was done in the previous years.

So we just wanted to alert you to that because Nicole will probably be sending you through an email with some instructions and questions for you to think about prior to the in-person meeting.

So if anybody has any questions about the upcoming process or steps, you should also have received travel details from our travel people here at NQF. No questions?

All right, well, thank you very much for your participation on the call today. Let us know if you have any questions moving forward and we're looking forward to seeing you at the end of February.

Suzanne Heurtin-Roberts: All right, thank you. Bye-bye.

Lorna Lynn: Thanks so much.

END