NATIONAL QUALITY FORUM

Moderator: Benita Kornegay Henry
May 2, 2019
3:08 pm CT

Rose Baez: Hi, this is Rose Baez.

Connie Anderson: Rose, hi, this is Connie. I just joined too. I think we're the only two on.

Rose Baez: Okay, thanks.

Woman: Hello. I just want to make sure that I could be heard.

Operator: Sub-conferencing is no longer active.

Woman: Hello.

Woman: Hi.

Woman: Hello.

Woman: Hello.

Man: Hi.

Woman: I think there are a bunch of us here but ...

Woman: I know there's a bunch of us here, but the (office) is not active.

Woman: Did anyone else hear the recording that the session was no longer active?

Man: Yes.

Woman: Yes.

Woman: Yes.

Woman: Yes.

Woman: Yes.

Woman: Yes.

Woman: I’m not sure if we should hang up and dial back in or what?

Woman: (Unintelligible).

Woman: (Unintelligible) dial back in?

Madison Jung: Hi, everyone.

Man: (Unintelligible).

Woman: Hello.

Madison Jung: This is Madison with NQF. I'll turn it over to Ashlie to open up the Web meeting. Thank you for your patience today.

Woman: Thank you.

Ashlie Wilbon: Good afternoon, everyone. This is Ashlie Wilbon. I'm a Senior Director here at NQF and I'm helping to lead this effort. I am a newer member to the project team. I think this is the first webinar that I've been a part of with the committee. So, I'm excited to be here and happy to be able to work with all of you.

Just a brief introduction of myself, I've been with NQF for about 10 years now working on various efforts here, a lot of time spent on the cost and resource use measurement efforts as well as helping with measure maintenance and the scientific methods panel team.

So, happy to be a part of this discussion and looking forward to hearing your insights on what we've got prepared for today. I will -- I would also like to welcome my team members. I'm joined by Jean-Luc, Madison and Navya here who have been a great -- done a great job of getting me up to speed on their hard work to-date.

So, with that, I'm going to go ahead and hand it over to Navya and the team to say a quick hello and get everyone up to speed on the agenda, our objectives and plan for today. And also, actually, before I do that, why don't I allow Eddie and Rose to provide a welcome to the committee as well?

Eddie Machado: Go ahead, Rose, if you want to start.

Rose Baez: Oh, sure. Hi, good afternoon, everybody. So, I'm looking forward to another productive meeting and lively discussion. From the final environmental scan report that NQF sent earlier today, you know, it's really apparent just how important this work that we're doing is and helping close the gap in communication and feedback between those to develop, evaluate and select measures and those to implement measures. So, I'm looking forward to today's discussion.

Eddie Machado: Good afternoon, everyone. This is Eddie. Likewise, I just want to welcome everyone back. You know, the committee has done some very good work to-date and I think we have a full agenda over the next few meetings to really further explore the existing options that NQF uses to capture feedback on the measures and really see if we can provide some input on the part of the CDP process on the use and usability criteria. So, looking forward to a lively discussion.

Navya Kumar: All right. Thank you, everyone. This is Navya. If committee members could please place yourselves on mute if you're not speaking.

Man: (Unintelligible).

Woman: Oh, let's see if we can mute your line.

Navya Kumar: Okay. Or we could try to mute.

Woman: Sorry. Hello, can everyone hear us? If you are speaking right now and your line is not muted, please mute it.

Man: (Unintelligible).

Madison Jung: Okay, Okay. I think we've muted it. Go ahead.

Navya Kumar: Thank you, Madison. Like I was saying, for those who have not placed yourself on mute, please do so if you're not speaking. And please announce yourself if you are going to be speaking and raise your hand if you wish to be called on. And feel free to use the chat box to send messages to the staff and co-chairs and we'll do our best to bring them to our attention to the group and co-chairs as needed. And for our public participants, please note that there will be a moment later on in the Web meeting for an opportunity to provide any comments.

And with that, we shall begin. So, our agenda for today is to provide an overview of NQF processes and evaluation criteria as well as specific to the use and usability criteria and submission form.

Hopefully, we'll have enough time to also go over the other feedback considered for the evaluation use and usability. If not, we'll continue for our Part 2 on May 7th and go over that then along with the challenges and gaps. After that, we'll go -- we'll have an opportunity for public comments and provide next steps.

So, for this Meeting 3 and 4, Part 1 and 2, we'll provide an overview of the use and usability criteria and as well as discuss current feedback channels, target audiences and inputs to the evaluation of use and usability. We hope to engage the committee for gaps and challenges in the current mechanisms as well as develop recommendations to improve the solicitation and collection of feedback.

And so, with that, we'll begin our roll call for today. So, Rose and Eddie, you are here. Connie Anderson?

Connie Anderson: I'm here.

Navya Kumar: Thank you. Robert Centor?

Woman: He is on the line but I believe he'll be joining us.

Navya Kumar: Okay, thank you. Elvia Chavarria?

Elvia Chavarria: I'm here.

Navya Kumar: Thank you. Dan Culica?

Dan Culica: Here.

Navya Kumar: Thank you. Melody Danko Holsomback?

Melody Danko Holsomback: I'm here.

Navya Kumar: Thank you. Anne Deutsch?

Anne Deutsch: Here.

Navya Kumar: Thank you. Tricia Elliott?

Tricia Elliott: I'm here.

Navya Kumar: Thank you. Lee Fleisher?

Lee Fleisher: I'm here.

Navya Kumar: Thank you. Mark Huang?

Mark Huang: Here.

Navya Kumar: Thank you. Joseph Kunisch? Joseph, are you on mute? Okay, Claire Noel-Miller?

Claire Noel-Miller: Yes, I'm here.

Navya Kumar: Thank you. Ekta Punwani?

Ekta Punwani: I am here, thank you.

Navya Kumar: Thank you. Koryn Rubin? Beth Rubinstein?

Elizabeth Rubinstein: I'm here.

Navya Kumar: Thank you. Sue Sheridan? Jill Shuemaker?

Jill Shuemaker: I'm here.

Navya Kumar: Thank you. Heather Smith?

Heather Smith: Here.

Navya Kumar: Thank you. Deborah Struth? Sara Toomey?

Sara Toomey: I'm here.

Navya Kumar: Thank you. And did anyone else join the line? Joe Kunisch, Koryn Rubin or Sue Sheridan? All right. Well, thank you. I will pass the presentation to my colleague, Madison, to go over the project background.

Madison Jung: Okay, great. Great, thank you, Navya. So, just to remind everyone where we are on this project, so we -- as we mentioned earlier, you just received the final version of the environmental scan. Thank you, everyone, for your comments and review and thoughtfulness and helping us pull all of that together.

But we are now onto our next deliverable, which is the CDP use and usability document. The goal for -- the way this document fits into the overall project is that the environmental scan and this document will serve as the foundation for us to develop our options for the pilot of the measure feedback loop. And from those options, we will select a pilot and develop an implementation plan for that. So, that's just to get everyone a reminder of where we are and we're onto the next stage.

Again, another reminder of just the operational definitions of what we're working with. I wouldn't read these in detail but just to remind everybody, these are our objectives and this is the definitions that we are currently working with.

Great. So, then I will pass it over to my colleague, Jean-Luc to go over the CDP process and the NQF evaluation criteria. A quick note, we -- in case you weren't able to find the evaluation criteria on our Web site, we have added it to the meeting invitation just a few minutes ago. So, if you'd like to open that up and follow along, we'll be referencing that throughout the -- throughout the presentation.

Jean-Luc Tilly: Great, thanks, Madison. So, I'm just going to quickly review the criteria that NQF uses to evaluate measures during the Consensus Development Process. It is a process that's undergoing continuous revision I mean most recently. And as some of you may remember, We did a process improvement or Kaizen event in kind of in the middle 2017 that led to several changes, especially the new move to two cycles every year, which you, all -- which shows to reflect shortly.

So, these -- you know, the criteria that we use to evaluate measures are also refreshed pretty frequently and generally just about every year during the summer. So, as I said, you know, I mean, of course, NQF has been getting a little bit stricter over the years at raising the bar for enforcement. So, we'll just -- we'll go through all of that -- all of that.

So, this timeline you have here in front of you us intended really just to give you an overview of all the different steps in the process. And the first being that intend to submit process at green bar you think in the middle there, therefore we asked developers to submit the measure specifications and the measures testing for scientific accessibility. And if necessary, if the measure is a complex measure which is to say in an outcome measure, it will be reviewed by the scientific (unintelligible) (panel).

So, then in the next - the black bar there is the full measure submission deadline of about three months after intent to submit. That's where developers are adding in the evidence for the measure or the documentation of the feasibility of that measure, all of which -- of these, we'll go into shortly. But there's also the information that's a particular concern for us today that the feedback on the measure that makes up the bulk of these (unintelligible) evaluation.

So, then, you know, just a little time period for staff to review and the committees to take a crack in. There are 14 standing committees, you know, across a variety of different clinical topic areas like a cancer or surgery. They make their decisions, they refer that decision to the CSAC, which considers the standing committee's decision but then also the public comments that have been issued up to that point.

And then a final decision by that group, the Consensus Standards Approval Committee, the CSAC, lead to an endorsement decision for a measure or not, quickly followed by an appeals period where, you know, the third parties or others have an opportunity to appeal the decision, enter new evidence for consideration, et cetera.

So with that, I think we'll just dive right into the criteria. So, as I've said, the criteria have evolved quite a bit over time. You know, NQF is really trying to be in-tuned with the field. You know, we're constantly in dialog just about every month with measure developers, with CMS, with NQF's own membership that covers just about every stakeholder group.

Our goal is, you know, to keep really what's changing a measurement, how many to adopt our criteria, how can we be responsive to those changes, you know, particularly in new frontiers like eCQM.

So, the five major evaluation criteria are -- so the first is it's important to measure and report. And really, we think of this as being two different things. There needs to be evidence that the -- of a relationship between if you're measuring a process of care, evidence of relationship between that process and between a positive health outcome.

You know, conversely, if you're measuring an outcome, a real empirical relationship back to a process or the health care intervention that establishes that that outcome is under the clinician's control.

And then the second part of importance to measure report is an opportunity for improvement. So, you need data to establish that there is a performance gap that a performance on the measure could be substantially improved.

The next is a scientific accessibility of the measure. So, here, we're looking for evidence of the measures producing consistent results, which will cover liability, and that these results correctly reflected quality of care being provided.

So, generally, validity testing is looking at - but the data elements of the measure and the measure results and making sure that the scores reflect differences in care as compared usually to a kind of a separate authoritative scores.

So, you know, sometimes, that's face validity and extra panel is the (associated source) or sometimes when we're thinking more of empirical validity testing, and that's a comparison with maybe another health care performance measure or some other differentiation in quality that represents a gold standard.

So, then, feasibility is pretty simple really just looking at the degree of burden that would be imposed by implementing the performance measure. So, you know, some of the questions we're asking there are, are the data elements already been generated or is it part of care delivery, you know, like is the measure using lab work, and are these data elements available in alternative sources.

And then finally, usability and use, so, and this one we'll go into really quite a bit of detail so I'll spend a lot of time in here. But, you know, it's suffice it just to say, the (instruction) usability and use is looking at whether the performance measure is actually being used or, you know, alternatively, does it have a kind of an incredible plan for future use, and whether the measure is driving improvement in the actual health care setting, you know, with minimal unintended consequences.

So, what you'll -- what you'll -- of course, what (you're thinking about) I'm sure as you hear all that is that really, across just about every criterion feedback comes into play so -- and important for the measure results (are in forming) the performance gap question, the opportunity for improvement, you know, and feasibility, they're all questions around various stage two implementation (unintelligible) and feedback.

And, of course, usability and use is explicitly asking for feedback for information about unintended consequences, information about how the performance measure results are being used by clinicians and whether those results are being used in an accountability application.

So, when the standing committee is evaluating usability and use, they're looking at feedback from really a variety of sources, all of which we will touch on if not today then later on May 7th.

So, the -- all these sources which NQF is compiling into a single document for the NQF to take a look at. So, we're pulling feedback from what the -- what the developers (unintelligible) but then also collecting feedback from the different public comment periods, the measure applications partnership, workgroups, a little bit of a (separate) (unintelligible) which we will explain,

and from NQF's feedback tool just primarily oriented towards, you know, of course, the public but then also NQF membership.

So, now, just to summarize the, you know, the evaluation criteria, we're looking at our assessment and updated on an annual basis as we said to kind of keep up with the quality measurement plans.

You know, these criteria are reflected in that document that we sent around earlier. And the changes in the criteria are reflected back in the measure submission form. So, our goal is to make sure that we're asking developers to provide information that were directly using them in the evaluation of that measure. And, you know, not imposing an undue burden on the measure developers themselves in terms of putting together the submission.

And so, with that, we're going to dive a little more deeply into the use and usability criteria specifically I'll be passing off to Ashlie.

Ashlie Wilbon: Thanks, Jean-Luc. I think this might be a good opportunity to just pause. I know Jean-Luc went through quite a few kinds of major concepts and our criteria, the various inputs that we've received into the process. And so, I just wanted to pause here and see if folks had any kind of baseline questions about any of the content that Jean-Luc covered before we kind of dive into some of the more detailed information around use and usability and some of those discussion questions. Any questions that people may have before we go a little bit deeper?

Elvia Chavarria: Hello. This is Elvia. I do have a question with the PCPI. And I know that before when we -- when we fill out the use and usability section, I know that I believe that usability is not a must-pass criteria.

And I'm sorry if I missed it within Jean-Luc's talking right now. I'm wondering whether there is a move then to make this a must-pass criteria.

Ashlie Wilbon: So, thanks for that question. We're actually going to talk in detail about the use and usability criteria coming up here really shortly. And that is a topic of the session that we'll raise for the committee to discuss about whether or not that should continue.

Actually right now, only the use is a must-pass for maintenance measures. And so, we'll kind of dive into that a little bit more. But I think that specific question has come up very frequently and I think as we kind of consider the importance of this particular criteria and really making the discussion about use and usability meaningful in the context of the other criteria, whether or not it makes sense to reconsider whether that is a must-pass criteria. I think it's certainly a question from the table. So, we'll certainly talk about that coming up here shortly.

Elvia Chavarria: Sounds good, thanks.

Elizabeth Rubinstein: Yes. This is Beth Rubinstein. Can you hear me? Hello.

Ashlie Wilbon: Hi.

Elizabeth Rubinstein: Can you hear me?

Ashlie Wilbon: Yes.

Elizabeth Rubinstein: Okay. As a patient of (unintelligible) there's maybe a language convention, which just even goes back to (unintelligible). And we just

mentioned again the term stakeholder. And I find that stakeholders with different measurements are different groups.

We don't have universal stakeholders throughout, so when we speak about stakeholders, do we ever define our (unintelligible) stakeholders or is this a universal convention to (be accepted) but when we look at particular measures, the stakeholders are different especially when you want patient feedback or public feedback from patients?

Ashlie Wilbon: Right. I think that's a very great point. I think in the context of the report, the environmental scan report, we identified I think four different stakeholder groups, if you will, that may be where you met kind of active participants in the feedback loop at different points in the process.

So, we talked about patients, we talked about the measure developers, we talked about those who are being measured, so that may be providers or what have you and then also kind of the administrators or kind of measure implementers or measure users as another stakeholder group.

But I think you bring up a good point. I think we use that term very loosely often to refer the kind of everyone who may be interested in engaging around either the use or implementation of the measure. So, that's something that I think we can be a little more specific about going forward, thank you.

Elizabeth Rubinstein: Thank you.

Ashlie Wilbon: Any other comments or questions before we move forward into the use and usability section? Okay. So, I'm finding my place here. So, we're starting at Slide 19 and to Slide 20.

So, usability and use, so this definition that you see here on the screen reflects our current definition for this criterion, which is the extent to which potential audiences, which may include, to the question that was just raised about stakeholders, so consumers, purchasers, providers or policymakers, and which would include all of those groups, are using or could use performance results for both accountability and performance improvement to achieve the goal of high-quality efficient health care for individuals and populations. So, that is the definition that we're currently working under and we'll continue to discuss how that definition is reflected in our submission form and discussions with the committee.

So, we wanted to give you a little bit of background that might help support some of the discussions and recommendations for today. So, this effort is not the first that we've had, focusing on usability and use. Going back to 2011 and even prior to that, usability has been one of the four core criteria that we had in place for several years.

Slide 20 really just illustrate some of the major reports and kind of changes or activities we've had around this criteria over the last several years. So, usability, going back to 2011 before we updated the definition, you can see the definition that we had driving those evaluations back in 2011 was the extent to which intended audiences are kind of same stakeholder group there, round of consumers, purchasers, providers and policymakers can understand the results of the performance measure and find them useful for decision-making, so, again, a slightly different definition there.

And following that definition, we convened a multi-stakeholder group in 2012, which resulted in a report, a usability report that is linked in this presentation. So, I would encourage you to take a look at if you haven't already done so.

But there were quite a bit of thinking around -- done around this criteria and some of the principles that kind of drove their thinking and recommendations about how this criteria should be evaluated, what are the various aspects that should be considered when evaluating the usability of a measure, and they also put forth some recommendations.

And part of the recommendations that were put forward as part of that report was to split out feasibility criteria to two sub-criteria, which is use and usability, which is what we have standing today.

And so, the official split was shortly following that report of 2012 and that was reflected as 2013. And then 2017 was when we transitioned to the must-pass for use for maintenance measures. So, again, a bit of a timeline there to give you a sense of some of the work we've done over time and this criteria evolving with the input of various stakeholders -- various inputs from -- I'm checking myself on use of stakeholders. So, inputs from various groups that we've convened, that included patients, measure users, those being measured, so physicians and clinicians as well as health plans and other folks who implement measures included on those groups.

So, one of the things we wanted to do today is to just bring you back to some of the work that was done in 2012, which probably -- which was probably the last kind of focused effort on this criteria.

And several of those principles that were put forward I think are very relevant to the work we're doing today and looking at these feasibility criteria again. And so, we wanted to put those back in front of you again.

We bolded a few on the slide that we think are particularly relevant to this work to put in front of you to see whether or not you believe based on the context of our current quality measurement enterprise, NQF processes and the kind of temperature of our -- the measure development process and lifecycle today if you think these principles still hold true in terms of how we should be thinking about evaluating measures for usability and use.

So, I wouldn't read through them all in detail, but I will focus in on the ones that are bolded and then I'll hand it over to Rose and Eddie to open it up for discussion on some of these principles to see whether or not we think again that these are still relevant or whether or not there are additional principles that you think should be added to this that will guide your work going forward and/or additional recommendations for kind of foundational conceptual thinking for this work.

So, the two on this slide that we'd like to call your attention to around the bottom two here, to achieve maximal effects on quality health care over time. NQF-endorsed measures should be used in all applications for which they provide useful information.

So, again, just promoting the use of measures, where they -- where they can be used. Public disclosure, performance results not only is necessary for some type of selection for consumer choice but also ensure the accountability and provide external motivation for performance improvement.

On the next slide, here, measure developers may not be responsible for implementing performance measures for accountability or selection or quality improvement and may not have access to the required data or information about measure use.

The NQF criteria of importance to measure a report and scientifically acceptability of major properties to ensure that the measure is potentially useful for a variety of applications. And so, following that, the measure -- measure can be more or less useful -- a measure can be more or less useful to intended audiences, depending on the conditions of implementations for specific purposes.

And the NQF criterion for feasibility, particularly regarding the data required to implement measure performance also imposes usability, however, feasibility uses issues may be mitigated or the benefit of measuring performance may outweigh associated burden.

And I think one of the other kind of underlying principles of this -- of this that is reflected in the recommendations that I think will be I'm really interested to hear your feedback on is that the usability should be addressed -- usability and use should be addressed last in the hierarchy of the four major criteria because if the other criteria are met then the measure should be usable.

So, this assumption that if the committee have already deemed that the measure is important to measure, that it's scientifically acceptable and that it's feasible, that usability is not an automatic but there's an assumption that the measure having met those three criteria should also be usable. And so, that was the underlying kind of assumption to making the use and usability criteria (not a must pass) that because the other three are must-pass, the usability and uses is a bit of -- I don't want to say, given, but it is something that is assumed to be met at that point.

There - another recommendation that we wanted to focus your attention on is around the expectation for timeframe to achieve use and accountability

applications during public reporting, but that some flexibility should be allowed.

Again, this recommendation is really what has driven some of the changes in the submission form around developers providing the rationale for whether -- particularly around maintenance measures that there is a plan for the measure to be used within certain timeframes. And that also another recommendation around the benefits of measurement in terms of facilitating approvement should outweigh evidence of unintended negative consequences.

And with that, I am going to hand it over to Rose and Eddie to guide us through a bit of discussion around these principles as well the recommendation that were put forward in this prior work to see their relevancy for the -- this current effort.

Eddie Machado: Okay, thank you, Ashlie, for covering that. So, it's quite a bit, you know, to digest all at once. So, I would ask for folks interested, and I'd say it's just to use the Raise Your Hand function in the chat and we'll use that ads our guide. So, I think Lee has his hand up. Maybe you can go ahead.

Lee Fleisher: Sure, and thank you. So, the interesting thing is the relationship between this and the public reporting, which I forget where that - is that inside this, the requirement for public reporting or separate?

Ashlie Wilbon: Public reporting was I think one of the -- I think if I -- are you referring to the discussion on the principles or where an evaluation criteria, the submission form that is specifically addressed?

Lee Fleisher: Where in the submission form the requirement for maintenance measure to be publicly reported?

Ashlie Wilbon: Yes. I believe it is in the submission form. I have to pull up the specific language that we used there. But there is some language around that use, yes.

Lee Fleisher: And it's about use.

Ashlie Wilbon: Yes.

Lee Fleisher: So, this became a major debate at the surgery steering committee of which I'm co-chair, and that the Society of Thoracic Surgeons had endorsed sub-measures, basically components of a composite. And the ability to -- and they wanted the component to maintain their endorsement.

It becomes a really interesting question in the use and usability of measures and that how do you display them and what people can absorb in guiding actions is quite complicated. And in fact, they tried to simplify as a way of saying, you know, above, below or at the national norms for the composite.

But in the individual measures, they thought it would be -- some of them they showed, some of them they didn't. And I think in some of these principles around usability, we have to ask the question of, you know, is it usability from just a public reporting stage, is it usability from the providers themselves of understanding what aspect of the metric helps drive change because they're very different?

So, I just think as we think through this usability -- use and usability, we have to ask, is it for everyone or for subcomponents, and how easy is it to -- not everything is fully understandable at the individual measure, particularly as it's part of the composite.

Eddie Machado: So, this is Eddie, thank you, Lee. That's a really good point. And I guess I just want to probe that a little bit further. Are you -- you know, are you generally sort of getting at the idea that usability is really predicated on the context in which the measure is being used or applied? Is that more or less what you're trying to get at or is it more specific to the component -- the portion of the measure that user ultimately is determining or making a judgment call on whether to use or display for use by the -- by the (audience)?

Lee Fleisher: And actually, that's why I asked about where -- I forgot where in the criteria this is important. But, you know, we've had these components for -- we've had endorsed measures for two decades to some extent.

And the ability of the public to understand varies greatly. And some of this depends on the complexity of what we're measuring. So, as we -- and particularly for composites and if we want to look at the issue, endorsing both the component and the composite in total, I think we need somewhere in this language that ability of, is this broadly defined of all stakeholders, some stakeholders, how do you best -- what is minimum criteria? Hopefully, I'm making myself clear. You know, I'm trying to -- it's actually about two public reporting and use.

Eddie Machado: Yes. So, would you, I guess, along those lines, would you call into question the language on the third bullet at the slide that's currently being displayed and specifically on the second line, the portion that says, "Use in all applications." Would that be a potential place where that could be addressed do you think?

Lee Fleisher: I think so, that's right. And that for which they provide useful information, I think that, you know, we need to weigh the utility versus the burden of

explaining. If it still can be used, for example, a component of some of the measures, the providers may find it very useful ...

Woman: We have a (unintelligible) to talk, right?

Lee Fleisher: ... to drive care.

Man: Sure.

Woman: Okay, I'll let her know. She (unintelligible).

Man: Okay. Am I in trouble?

Woman: Yes.

Ashlie Wilbon: Hi. This is Ashlie from NQF. I'm just going to ask if you could just mute your line if you're not speaking. Thanks. Is Elisa Munthali there? Are you there, Elisa? I know she was trying to get a word in. Are you there?

Elisa Munthali: I am. Sorry about that.

Ashlie Wilbon: Okay.

Elisa Munthali: So, hi, Lee, this is Elisa.

Lee Fleisher: Of course, you are there to answer my question.

Elisa Munthali: That's okay. It is an interesting question about what transparency means. I think the guidance that we have really does point to public reporting at that

highest level of transparency. But it's not prescriptive about other ways in which measures can be transparent.

With regards to the surgery review of those component measures, because the components were submitted as individual measures outside of the composite, they were also subject to review of -- a review against NQF's entire criteria. So, that's why it was challenging.

I think, you know, this is a discussion that the CSAC, our Consensus Standards Approval Committee, started, you know, raising last week. They're talking about, you know, whether or not, you know, debating big use and usability criteria, whether or not the guidance is clear, we may be pulling together technical expert panels to look at that guidance to make it a little less confusing and clearer for developers and other stakeholders.

Lee Fleisher: Thank you. And I do think the accountability and transparency and use and usability are related.

Elisa Munthali: I think they are as well, and that's something where the CSAC also noted. We walked them though the language and the guidance then I think it could be -- you know, the language could be expressed a lot clearer.

Lee Fleisher: Thank you. That's great.

Ashlie Wilbon: Hi. This is Ashlie. I would just add that obviously, the efforts of CSAC and this committee are complimentary. So, to the extent that this group has recommendations, they would ultimately -- for the criteria, they would ultimately be considered by the feedback.

And I think having input from other bodies like this committee would be helpful for them and their consideration as well. So, I just wanted to add that in as well.

Eddie Machado: Do we have other questions? I see in the chat box, Elizabeth, you've written a couple of notes I think in relation to Lee's comments. Do you want to comment further or no?

Elizabeth Rubinstein: Well, I believe Lee was touching upon a few wishes there from a layperson patient advocate to your point, identifying the end-use of a mission that placing to the pre-design and pre-testing for public patient caregiver consumption.

And the ending is this versus professional quality improvement measures or measures, you know, to what leads to our medical professional, the designing factors are different. So, one size does not fit all.

And clearly defining the use and usability needs to be broken down through different audiences because they're designed differently.

Eddie Machado: It's a valid point. I mean it ready does vary depending upon the audience, the way in which it's being used or the program, and there's a lot of variance there so.

Rose Baez: I see Mark, you have your hand raised. Do you want to ask a question?

Mark Huang: Yes. You know, in terms of just following up on some of the other comments regarding public reporting and sort of going back to that one third bullet point, you know, to achieve the maximal effect, measure should be using all applications especially providing this information.

I think, you know, part of the struggle especially with sub-specialty providers is that they are going to report on a lot of measures that are more general or not really -- they're not really representative of what that practitioner does on a daily basis. They're just sort of reporting just for the sake of reporting for quality payment, quality program kind of things.

And that's where the danger falls in. So, if these people aren't performing as well and say, I don't know, a measure that might be more relevant in primary care such as either hypertension or it could be retinopathy screening or something like that where, you know, that's not really the focus of that practice.

You know, if they come across these issues, they may manage or refer up for management and that's where the danger lies and sort of this public reporting is just like as people have mentioned, it's taken out of context, say, "Oh, this person is a lousy provider." And, you know, it's sort of -- there was a lot of explaining to do on the backend in terms of, "Okay, well, yes, we report these measures where we need to report these measures."

And there aren't not measures that really accurately reflect what we're doing on a day to day basis. And I think that's sort of the danger in sort of that reporting -- public reporting just for the sake of just reporting something.

Rose Baez: You bring up some valid points there. And I'm wondering as we look at these principles, is there a way to address or add additional principles to get to your point?

Mark Huang: I almost feel like it's -- you know, obviously, it's getting misused, but, you know, the relevancy to the practice has got to be factored into sort of this sort

of reporting. You know, we sort of -- you know, if a -- we're just going to take, for instance, some other -- you know sometime in physical (unintelligible) patient, you know, there's not as much (unintelligible) that care for specific population they are seeing, say, brain injuries or head injuries, spinal cord injuries or stroke or (an implication) or it's more of a relocation focus aspect of it may not be the (current) treatment or prevention.

We still obviously are in-tuned with secondary prevention, complication, things like that. But, you know, there is a measure that really reflects that in day-to-day practice. So, where you left with, you're back to generic obesity, hypertension, smoking cessation, you know, some other things and falls as somewhat of a relevance, you know, that (unintelligible) back to, you know, sure, we do well on all risk assessments and cognitive assessments.

We all do quite as well on hypertension screening and management improvement because we're not really seeing them for the improvement in blood pressure. We may note that they have high blood pressure and oftentimes may coordinate with a primary care physician (with a tie) we're not actually managing that. So, this is back to, you know, the relevancy or does that really reflect what that provider's practice really is when you're talking about reporting?

Rose Baez: Sure, because that can lead to into some unintended negative consequences ...

Mark Huang: Correct.

Rose Baez: ... (unintelligible), right?

Mark Huang: Right.

Rose Baez: Well, thank you for that.

Eddie Machado: All right.

Rose Baez: Go ahead, please.

Eddie Machado: Sorry. I think we see Dan has raised his hand. Do you want to go ahead?

Dan Culica: Sure. I was just thinking in terms of reporting regardless whether they did public or not maybe as a principle or a guiding principle. I think that usability is something that comes in more as before the fact why they're used, it's more after the fact.

So, I'm thinking if whether some kind of a scientific method like a correlation analysis or something on that line between the usability and the use when the reporting gets done can influence the sort of the conclusions of the reporting.

Eddie Machado: Okay. Do other have thoughts on that? I mean that's a different way of maybe approaching this.

Jill Shuemaker: Hi. This is Jill Shuemaker. And I just wanted to address the public reporting and how it impacts physician who are maybe doing specialist care. So, I think we should keep in mind that the goal is to have a variety of measures, so clinicians have the ability to choose measures that are relevant to their practice.

And I understand that in some group report, large group reporting situations, multispecialty clinics for regulatory reporting, sometimes you have to report on measures that the outcome reflects the group and not necessarily one specialist.

But I think we just need to just remember and not get bogged down on the fact that a physician who is specialist may be having to report on a measure that's not relevant to him because hopefully, we have enough measures that he can choose the measures that are relevant to him and that if this measure isn't relevant then he wouldn't choose it and so therefore he wouldn't run the risk of it being publicly reported and negatively impacting the practice.

Man: So, I'm going to counter on that. I mean I do agree just in general with, you know, obviously measures like hypertension, obesity, smoking cessation, they're really relevant to all providers.

So, you know, I think we have an understanding in terms of obviously sort of measures across all specialties that are important. But I think the one downside we have, our current limitations like, so for instance, (I'm going to use) our organization and many in our specialty that have this problem is that we actually don't have other choices especially if we're going to do certain styles of reporting for quality.

So, for instance, if we choose the electronic health record submission, we're limited literally to about 35 measures, of which we struggle every year with which ones of the best of the least relevance.

And so, that's the problem here, is yes, that across the board, there are many measures that are available, but it depends upon a reporting mechanism in terms of what are the measures available for you to choose from. So, that's where it becomes challenging.

And I think some of the quality payment program reporting mechanisms are going to be more flexible for the upcoming years, but then now, we're in a

reporting burden issue which is you're going to say a report, a couple of measures through a data registry and maybe a couple of registers through group reporting on the CMS Web site and maybe a couple through EHR, electronic health record, means of submission.

So, that is very burdensome to actually choose in multiple. So, now, this is a separate issue, here's actually a feasibility of actually reporting across multiple measures as opposed to the feasibility reporting in one measure.

So, you're just reporting for, say, quality scoring for the CMS quality payment program, it becomes very burdensome to choose more than one means of reporting. And, you know, in the past, they've actually limited you trying to go through one means and that can automatically actually exclude many, many measures for me to use.

So, that's sort of -- that's sort of something outside the scope of this conversation, but just that it gives you a sort of a landscape on what the realities of providers are faced with when they're trying to do the sort of reporting (on data).

Connie Anderson: This is Connie Anderson. And I could share the renal standing committee. And I echo what you've just said entirely. And I'm really concerned about the benefit to measure the principle, the benefits of measurement in terms of (unintelligible) improvement should outweigh the evidence of unintended negative consequences.

And many of the measures that come out were more facility-driven than clinician-driven through the ESRD, end-stage renal disease facilities. But I have some real concerns about how measures, the use and usability may be

targeted at the wrong entity and there were tremendous unintended negative consequences.

And there is -- and that whether it's publicly reported through -- we have five-star system or if it's publicly -- if it's publicly reported through the QIP, which us our Quality Improvement Performance system, which is paid -- and it's pay for performance but it's certainly a penalty system.

So, I think that I'm real concerned about -- I think we need to do more with the principle about the unintended consequences that define it better because I think if people don't understand the process when they're making decisions about whether or not these are measured through these publicly reported and what's available for the clinician or to the facility in our case. I think we really do harm and we're not facilitating and so that we're actually causing unintended consequences.

Melody Danko Holsomback: And this is Melody Danko Holsomback. I, you know, I agree with that and, you know, some of the other things that are -- along with the different types of reporting, we're an ACO so we group report, but with that we're in conflict with the rest of our system under what the benchmark should be because we do Web interface reporting and dig into the charts and do it manually for the most part.

And then we have HEDIS measures that a lot of times are just hold on our registry-type report and they have different values. So, our providers, you know, they think their scores have to be, you know, for the exact same measure for two different -- for two different payments might be our MA plan who has paid through Medicare as well as our Medicare MSSP, the same exact measure but there are two different expectations for what they have to reach.

So, that's another aspect of this as well, is getting everything consistent across the board. And if we're pulling them electronically, make that electronic pull that -- so that it's pulling the same type of data and the weight that you would get if you're doing it manually to get those numbers because it's very confusing and conflicting.

Sara Toomey: Hi, this is Sara Toomey. There is something -- there is another somewhat related issue. You know, I'm a pediatrician. And I think pediatric has other issues around this in the sense that A, they suffer from what was described as subspecialties that there not being as many relevant measures, in particular in our pediatric subspecialty but even more generally.

And secondly, there is not as much public reporting, so for measures developers, it's often extraordinarily difficult to be able to know how one measure is being used because there are no -- there are no real public ways of tallying the use. So, I think that's something else that we need to make sure we consider.

Eddie Machado: This is Eddie. This has been great discussion. I just want to check in with maybe Ashlie and the staff. I think we could -- we could probably keep going on this topic because obviously, we have very strong feelings about the issue of transparency and, you know, what the current landscape of measures looks like and the availability of relevant ones. But I'm not sure if you want us to keep going on this topic or whether you want to keep moving through the slides at this point.

Ashlie Wilbon: Yes, thanks, Eddie. Yes, I think we should move forward because I think a lot of the maybe specific issues will be highlighted within the context of the criteria. And I think maybe having those discussions, I think we have enough

here to know, maybe even propose a few additional principles that we made great pull back and forth in front of you at some point. But I think moving forward to talking about the specific criteria within the use and usability criteria of what we're asking for will help kind of focus the conversation a little bit more.

Jean-Luc Tilly: Okay, great. So, this is Jean-Luc. I think we can take it back up on Slide 25, the use and usability criteria domain. So, just I mean the first use is pretty conceptually straightforward and we had our conversation about this. But, you know, accountability and transparency really, we're looking at, you know, is the measure used in accountability application after three years or for local reporting program after six.

Feedback (unintelligible) measuring others is really just kind of prompting developers to tell us about how the measure performance results are being used, what kind of assistance is being given in terms of the interpretation of those results, you know, especially the clinicians and others who are being evaluated, you know, then summarizing the feedback from the same clinicians and eventually describing how the feedback, you know, was used to revise the specifications if applicable.

Usability is asking about two different things (unintelligible) documented improvements since the initial endorsement, and, you know, do the findings that are unexpected which is to say the unintended consequences or unexpected benefits, you know, how do those balance that?

So, on the next slide, just a little bit of information about the kinds of information we're collecting on the measures submission form from developers. So, in the 600 (some odd) submissions since December 2016, I

mean really, in terms of (themes), you know, the majority of the fields are basically blank or not applicable.

Very often, developers, you know, as we said, coming again, you know, just struggle to take in feedback, struggle to find a good way to loop that feedback back to those who are being -- who are being measured or otherwise, is not involved in the implementation of the measure. And so, don't really have an opportunity to immediately participate in a way that would give them anything to really (say).

When they do (unintelligible) (say), you know, many developers give us an inclination of their processes to give feedback, but less information generally on the feedback itself, you know, although we do get a fair amount of information about maybe the kind of clarification they've offered.

And then when we're looking at really at different kinds of feedback with just a little bit of information there, unintended consequences and implementation burden and how those are being addressed.

You know, one example might be, you know, unintended consequences of (unintelligible) under reporting of serious adverse events and then how they're kind of dealing with that in the measure construction.

So, now, just to go through some of these -- the specific criteria themselves and how they're written out, the use of current and planned uses, the measure, just, you know, as I said, I was to talk about making sure the performances are also used in really a form of accountability application within three years, publicly reported at six years after that initial endorsement, which is, and this is a fairly recent change now and must-pass for maintenance measures.

So, some of the key questions that we're looking at, you know, so at the time of the initial submission for a new measure, you know, what does the developer's plan look like for an implementation and accountability program?

You know, is it a credible plan, which we think of as, you know, including a specific program, a purpose and intended audience, a timeline for the implementation, you know, if there is a question of local reporting or, you know, other issues of data aggregation or other kind of mechanism to review.

And so then, you know, if it's a maintenance measure then we're looking more, is the measure being used in an accountability application after three years, after six in a (unintelligible) reporting application?

If the answers are no then, you know, what are some of those reasons that might be -- that the developers is using to justify that? (Unintelligible), you know, are the different external factors, was there revision to the measure, maybe they kind of let it to be taken out of the program, was it in a program and (unintelligible) since then been removed, we see they're having sometimes some measures that are used in some federal -- at the federal accountability program.

And then, you know, asking the question again, is there credible plan to either get it back and just hold the reporting program or otherwise (unintelligible) there for the first time.

Madison Jung: So, I'll just add a little bit more context, this is Madison. So, this slide right here that we have up is pulled directly from our measure evaluation criteria guidance. So, this is to give you a sense of the kind of questions that from the NQF standpoint, our criteria is based on what we're looking for, so how we're thinking about asking about accountability and transparency.

These next two slides I'm going to go over will be kind of the flipside, so, you know, these are the questions we're looking for. How are we asking these questions with (unintelligible) submission (unintelligible)?

Jean-Luc Tilly: Yes, great, thanks. Yes, a great point. So, then on -- in our submission form, you'll see the sort of the table that we ask developers to fill out within - indicating, you know, asking to classify the intended use of the measure. You know, they just check all that apply and you know, of course, we can, if you have any questions about what those terms mean, we can define some of those for you.

But, you know, accountability applications, just to be clear, we think of as applications using performance that identifiable, accountable entities, to make judgments and decisions, that have consequences for their performance. You know, it could be a different payment rewards recognition even or some other kind of selection keys.

And then public reporting is making comparisons - well making those performance results available, freely available to the public at large. Again, about identifiable, accountable entities, and then, you know, where possible making sure those are in such a format, such that comparison between different entities are possible

To the use in the submission form, you know, we've kind of gone over some of this already, but use is specifically have how we're asking those questions. So we'll just leave that up for a couple of seconds, then we can turn to the questions of the current state. So…

(Ed): So (Luke), this is (Ed), I know (Sarah)'s had her hand up. (Sarah), do you have a comment or a question?

(Sarah): Yes. I guess a little bit of both. I mean, I have to say making - for pediatrics depending on what you mean by both accountability and public reporting, it is an extraordinarily high bar given that there is no federal level sort of mandated quality measure reporting at this time.

And so I'm just interested to know whether or not - well so in other words, I think, I'll give an example of one of the measures that I'm the steward of, child age caps. I, you know, I know that it's being used in some hospitals, you know, insurance contracts and that sort of stuff by - but that's kind of because I know people, there's no way to know that in a public manner. And I know there are hospitals that are putting child age caps results out on their Web sites and things like that.

But once again, there, currently, there is no public reporting that occurs. There's no way for there to be any higher-level sort of reporting that happens.

So how you guys thinking through those types of situations where really it - the measure can't be used in the way that at least I'm understanding you're describing it?

Jean-Luc Tilly: Yes. Thank you. I think that ties really very precisely with the current state slide that I think you have it in front of you now. I mean, of course, first of all, is exactly what you said just now.

That yes, certainly it's our understanding that it's fairly common that a measure might be used in a program and the use of that measure might not get back to the developer. Especially in a way that allows the developer to, you

know, fill out the pretty specific requirements of the submission form. You know, that just the mere fact of it being used is somewhat not all that helpful.

And I mean, there are a few instances of developer's deciding not to submit fully endorsements because of that criteria or otherwise losing endorsements because of that criterion.

But even though this isn't exactly spelled out in the criterion as an option, there are some measures that have gone through without really tapping this criterion with a kind of justification or plan being advanced as an alternative to having actually met the criterion. And so those measures have retained endorsement.

So I would say indicates of the (Pease) measures mean that would be my understanding of what would be a pretty likely outcome. Is essentially the developer and the standing committees would both kind of recognize you know, where there might be some suitable accountability obligations. Public reporting just isn't really an option for most measures. And so they you know, that criterion might not apply or that justification would be made clear as part of the endorsement.

Ashlie, I don't know if you want to add anything to that.

Ashlie Wilbon: No, this is Ashlie. I think that's exactly right. And I think we actually have some discussion questions coming up here about whether or not I think to that last sub-bullet. We don't currently have a specific place in there for developers to kind of justify.

I think, you know, why it's not used in a public reporting function. I think that it's being provided may be in other spaces or as part of the discussion that the

committee will have during their evaluation process. But I think there is certainly some more thinking that we could do about how that criterion is evaluated and what the bar is for developers to meet for that particular sub-criterion.

Jean-Luc Tilly: Great. Thanks, Ashlie. So I think with that we just got one more slide before we get to those same discussion questions. So let's take care of that.

So the current safety - the use data that we have so far, you'll see most endorsed measures are in use and generally --I bet a simple majority are in use and in the kind of program that meets the standard -- and of course many of those new current use measures are new measures that haven't yet passed set three year or threshold for being called back for endorsement.

You know, however, it's certainly true that, you know, when you think it's a typical NQF endorsed measures may well not be used, or maybe used in a quality and improvement programs.

So it is something that we're thinking about in own former discussion questions in terms of, we think of how likely it is that a measured may struggle to find a good use case.

So with that, I think we could just turn over to the discussion questions and to Rose and Eddie

Ashlie Wilbon: And sorry, Jean-Luc. Just - this is Ashlie, just one more comment. I did also just want to add that public reporting is just one of the accountability applications that are - that we asked them to indicate. If you remember that slide with a table that (John - Luke) showed with all the different applications, the payment program, professional certification.

And so the questions that we ask in the submission form or the criteria around public reporting or another accountability application, so it's not all hanging on public reporting. I think the key is that it's being used in some applications. And I think public reporting obviously is the one that tends to be the most transparent. But other use and other types of accountability applications is also acceptable.

Ashlie Wilbon Okay. So let's just start with…

Man: Could you just - just quick question on that current use pie chart. Could you put some percentage of there? It's kind of hard to understand like you said, no current use is very small percentage, right? So - because it's that large blue? There's a lot of blues. It's hard to delineate which blue is to what.

Jean-Luc Tilly: Oh yes, yes, sorry. No, I, yes, - if I said that I actually misspoke. No, no current use. I don't remember the exact percentage, but that looks like about a third to me that have no current use. And I - many of those being new measures that we don't necessarily expect to have occurring use since we've only just earned endorsement.

The smaller blue sliver that you see is a quality improvement programs which is - so those are programs that are a use as a measure, but one that wouldn't meet that standard necessarily for accountability applications or of course for public reporting.

Ashlie Wilbon: Right. So the pie charts starts with no current use as the biggest blue slice and then it goes clockwise around down the list. That I think is the next iteration slice, we can include one that has percentages.

Jean-Luc Tilly: Okay. Yes. So let's turn to the discussion questions.

Rose Baez: Okay. Thank you. So just starting with this first discussion question, what expectations should there be for developer's system that use information, data submission form? Any thoughts around that?

Woman: To be frank, I'm not 100% sure I'm understanding the question. I mean, are you saying is the information that you're requesting on that prior slide appropriate or are you saying how much of that should we be expecting submissions to have? Maybe it's clear to others, I'm not 100% sure.

Ashlie Wilbon: So, this is Ashlie, I think (unintelligible) this is the question is really how much of those questions around use -- given the limitations that we've already discussed around developers having access to information around who's actually implementing their measures -- the question is kind of around, is the developer kind of the right audience to answer that question and what should the onus on the developer to be able to fulfill that answer on the submission form.

Man: And this is for initial measure submission, right?

Ashlie Wilbon: I think for both initial and maintenance measures, I think (unintelligible) in particular.

Man: Oh, maintenance, right, okay.

Ashlie Wilbon: Yes.

Man: Although it's interesting that many don't necessarily have the ability to implement, I still think they need, you know, developing a measure proposing

or maintaining that measure without having use information just seems, I just feel like that's still required element for them. They have to include something - well otherwise, why else are the developing the measure?

((Crosstalk))

Ashlie Wilbon: Yes. I also think that not the only the developer, that the user, especially if it's something that they're renewing, or slight or changing in any way - I bring up with a recent change in a CMS measure during an audit that affected ACO's. It was kind of modified midstream and had some downside effects for everyone.

So I think that expectation should be for developers and people that are planning to use it to have information on how they're going to use it on that submission form.

Connie Anderson: This is Connie. I agree. I think it's very important that the developers understand about the use information because otherwise as they construct the measure if they don't understand how it can be used the measure isn't going to be constructed correctly. And then it goes back to what that creates in those unintended consequences.\

So they don't implement the measures. But they absolutely should know how they would be used within the clinicians, or the facilities, or whatever because I think it does lead to unintended consequences if they don't know how to use the housing measure will be used.

(Sarah) This is (Sarah). So I think Connie brought up an interesting point, which I'm trying to remember we just submitted. I don't think you guys - well you kind

of asked that you could have a column for which it measure developers certainly should know how a measured could be used. That's definitely true.

I think measure developers should say what they know about how it's being used. Interesting that somebody reads - and I apologize, I just gave this to Melody, that others in the community should be saying how measures are being used. I don't think, I'm not sure measure developers always have access to that information. So I'm not sure that is on the - they'll be able to necessarily gather all of it.

But you guys do have a public report - a public sort of notice period, which I think is helpful. So if people do have comments, but I think I do agree that developers should give whatever information they have.

Ashlie Wilbon: Yes. I think when they're developing it, they should have some kind of a clause of this is the intended use. So if this is how it was developed for the intended use, so that if someone else is going out and they're going to use that measure, they have the context of which it was developed. So that those consequences don't happen if they're not going to be directly involved in creating it.

Eddie Machado: This is Eddie. I think that's a key nuance in terms of trying to thread the needle on what developers can and should do versus what they're realistically capable of knowing about how it is - actually the measures are actually being used in the field. Tricia, I know you have your hands up. Do you want to…?

Tricia Elliott: Yes, thanks. Kind of piggybacking on the prior comment, I think a specified the intended use and with the joint commission, we are a measure developer.

So typically we know how we're going to use the measure, but all of our measures are out in the public domain and we publish all of our specification. So there is a possibility the measure gets picked up and people use it for other uses as well. So we can only speak to why we developed the measure and what our intentions are.

The other piece that's a little bit difficult with the CQM's as we intend to use it in our program, either accreditation or certification. But if an EHR vendor doesn't pick it up and embed it in the software, then it may not get the use that we hoped for. So just wanting to offer that perspective.

Woman: I'd like to add another perspective to that. If a provider is in an organization that he finds himself fortunate to be able to choose what kind of quality measures he would use for his improvement activities that are outside of mandated reporting program, you'll often find within those settings that providers choosing measures that are relevant to their practice; that they don't use in a public reporting way.

So that is another use case or a developer not being aware of providers in some clinic settings that are utilizing the measures and find them very valuable. But you would not know that.

Eddie Machado: So this is Eddie that takes us, I think to the next question, which is if the developers are challenged to really identify that information and I think some would argue, you know, it would be a little tough for them to handle it. Are there other sources or stakeholders that we should be looking to deliver that information? And I think Mark, you have your hand up.

Mark Huang: Yeah, I was just, you know, there's that disconnect just because the measure developer may not have, you know, know all their use cases. You know, I

know this is not that I want to create more burden for reporting, but here was an opportunity. I don't know even say for example, in like the quality payment program, if we're doing quality measure reporting, is it just taking a moment, ask a few questions of each person who is reporting or group practice group?

You know, sort of like, okay, here are the - we know which measures they going to report on, but maybe here's an opportunity for them to provide feedback on those measures. Maybe with some set questions, you know, relevancy to the practice, you know, was it burdensome to report any suggestions for the developer or modifications?

This is an opportunity to get that feedback, you know, back for people who actually are using those measures and reporting on them in a more - and that could actually give information back to a measure developer. How potentially future measures can be crafted or are (unintelligible) gaps or modifications to existing measures.

So I mean, you - I think it's sort of closing the loop for measure developers would be very helpful in meeting them, giving them feedback from users. It's just the question of how can you automate that or encourage users to provide feedback to the original developers,

Ashlie Wilbon: Which is what I think we're kind of tasked with.

Rose Baez: Yes. And I think we might see this Rose you might be getting to this later in the presentation, but I'll ask maybe if (unintelligible) staff can comment. Is that part of the NQF form where they can provide that same type of feedback?

Madison Jung: Yes. This is Madison. So there is in the next actual section we're going to get into this feedback by those being measured by others and then there is an

opportunity to provide that. But I don't want to jump too far ahead, but I think a lot of the points that we're discussing as challenges here are also challenges more specifically for this next section.

Rose Baez: Okay. (Deborah), I see your hand is raised. Do you want to ask the question?

(Deborah): Well, actually just to make a comment, then just (unintelligible) with the Oncology Nursing Society and we have we are a QCDR vendor. And I think that I just wanted to clarify for the group within the quality payment program then we are permitted as a QCDR qualified clinical data registry to develop and test our own measures specifically for the MIPS program.

So they're custom QCDR measures and as a part of the requirement of participation, we then collect data on our measures and provide (SAMA) with performance data.

But we're also required to provide our users with data reports four times a year. And actually, our registry platform calculate data on measures from, subscribers, whether they're using the registry for performance improvement or accountability reporting every night at midnight.

So, you know, if the process of putting that data into the registry is a part of the workflow in a clinic, then a manager has - a decision maker, has near real-time data.

But you know, we continually then - we know who is using our measures. We have some idea of reliability and data quality as a result of that. So it's kind of a different application.

But one of the challenges for the QCDR, who - and we have several measures that we intend to bring forward and we come with that data and that experience.

So we're continually engaged in that feedback and use some guidance as far as technical expert panels to help with the development and testing measures also.

But one of the challenges is many providers aren't aware that QCDR measures are out there. They're not as easily accessible. I read gap analysis regular with (HIP out) measures in oncology and what the gaps are and very often QCDR measures aren't identified. So I just wanted to put in just a little bit of a different developer perspective.

Jean-Luc Tilly: Oh, that's interesting too. Just from a standpoint of, you know, this is the way a lot of gaps are going to be met is through specialties specific sort of data registries.

And absolutely, it's a great potential for subspecialties to be able to develop very relevant quality measures, you know, specific - especially the significant limitation is that a lot of times these measures that are being developed are only within those that are data registry participants or enrollees.

And many times the specialties struggle with exactly defining the measures that they want to measure. And so if - the challenge there as many times as they may not even have the resources or expertise to develop those measures. They may call in, obviously, a measure developer by contract to do those to - create those measures.

But again, they all live within really the QCDR itself and its limited sort of population. So granted a lot of times if it's specific to the specialty, but there's often many members of a specialty that may not even be aware of the QCDR or even participate. So it depends on the specialty. Like there's some specialties where there's very heavy use of the QCDR such as ophthalmology and cardiology.

But then there are others that are just barely scratching the surface of a QCDR. And it's - I think the challenge - it's a great area that needs to be developed further. Here is again a line in sort of question, is getting A, a developer hooked up with a data registry to help them craft the measures and then B help to refine them, you know, as time goes on?

It's a little bit easier to refine them when you're just sticking to your data registry participants, but then it becomes more - it become even more useful - then became a more broader applied measure that could be more easily accessible outside the data registry.

Eddie Machado: Okay.

Rose Baez: Any more comments on that topic before we move on? I did see Melody you have your hand raised.

Melody Danko Holsomback: That was for a previous comment. Thank you.

Rose Baez: Okay, there’s no other comments on that. Maybe we'll move to the third bullet here. Does it still make sense for use to be must pass for maintenance measures?

(Sarah): This is (Sarah). You know, I think once again, I can get to go to pediatrics, I think, you just have to be careful about how that gets defined and whether there are some justifications.

The other - I guess the other points that I guess I would like to make, in particular regards of also public reporting is, I think there's growing desire across the board for adults and pediatrics to become more parsimonious in the measures that we're reporting and that we're using. And in that context, you could imagine that there will be and that we might move towards distance by which people are either rotating the measures they're using and, or you know, trying to sort of really tailor a sort of a smaller number of measures for use.

And I just wonder whether or not - and so I guess the concern I have there is if all of a sudden this becomes a requirement in there that a bunch of measures no longer meet and so, therefore, they become unendorsed, are we then going to be hamstringing ourselves in the future in regards to having a body of rigorous measures that are being reviewed? Just as another sort of thought to consider.

Eddie Machado: It's a very good point and so I guess that leads us right into the last bullet, right? I mean, around, you know, having caveats or justification or a way to explain why there might be circumstances in which the way the criteria is currently written might not necessarily be applicable or relevant. I mean, how do folks feel about that? Mark, I see your hands raised.

Mark Huang: So I would think that should be an option. Just in cases and unusual circumstance, you know. We could say - when say the data registry in which they were involved with is somehow dissolved or, you know, we see data registries, vendors, things like that and migrate all the time. Vendor switches, alliances that occur are created and dissolved and then they get to look at their

measures again and then decide what's relevant, what's not or then measures being modified. That might be sort of like an explanation why it might not be utilized for and now, sort of in a limbo state.

Eddie Machado: Very good point. Melody I think.

Melody Danko Holsomback: Yes, I was going to just say something, you know, similar again when measures are picked up and I come back to this because it's kind of an area of we're working on now with our recent Webinar face audit.

Tobacco screening was altered and they used previous data that was not collected to justify increasing the benchmark and measuring a totally different measure. So justification on that. Things like that I think are needed because when your performance is dropped 3% if you're looking at the exact same measure from the previous year and they change the measure. It's dropping you, you know, 40 percentile.

There needs to be a way to justify that back and not get negative, you know, I guess negative scoring on your public reporting that it looks like you're severely deficient in something and it's not that you're severely, severely deficient, it's because something was changed that altered what you're actually collecting.

Eddie Machado: Yes. I think that's a very good point. I think the theme there is really, it sounds from both what you mentioned and what Mark mentioned, is that you know, there's particular circumstances that should be allowed to be explained or considered when evaluating this.

Melody Danko Holsomback: Correct.

Eddie Machado: Okay. Claire, I see you've raised your hand.

Claire Noel-Miller: Yes, just wanted to follow up on that and really agree that from, you know, from a consumer's perspective, if part of the reason why these measures, you know, quality measures are changing, is because there's been a change in what's being reported. Then it's really important for, you know, those measures that make it to the public reporting - it's really important for consumers to be able to have that transparency, to be able to understand the reason why they're changing in one direction or the other.

Eddie Machado: Thank you. Thank you. Any other comments from committee members on this topic or…

I guess Ashlie and others, I think we can probably move forward.

Madison Jung This is great. This is Madison. So this next slide is showing that second component of use that is also must pass for maintenance measure. We just reviewed that first component, the accountability and transparency. And this is the second component that we alluded to, the feedback by those being measured or others.

So this means that those being measured have given results and assistance in interpreting results. So those being measured and others have been given the opportunity to give feedback. The feedback has been considered by the developers.

So was there an opportunity for people to - or measuring users to, you know, ask questions, have their questions answered, does the developer have - has the developer proven that they've considered all these things? Maybe challenges or difficulties that measuring users have had?

So this again is just the questions that we ask based off of our measure evaluation criteria and guidance. So from maybe, the committee standpoint has the developer provided these points. These next few slides are what - how the questions are formatted in our measure submission form. So I'll just pause here and let you look over these bullets.

And this next slide is some examples of what we've gotten from our developers. So as you can see and as we mentioned (unintelligible) earlier, a lot of these - the feedback that we're getting is just summaries of the process, or that there has been no significant issues identified, or that just very high level, yes, these answer the question - these questions were answered. Yes, we've addressed their concerns.

This is what an example of one of the more detailed responses that we've received and a line by line of how they adjudicated them.

So again, to summarize - the summary of the feedback is provided is either not applicable or describe the process in which the feedback is collective and not the results. So, you know, all this information is useful to have, it’s not the most - I think we've heard from standing committees, it's not exactly what they're looking for and they would like the opportunity to hear more about the results.

And the flip side of this is, you know, often as you mentioned, developers do not have the ability to act on this information because they don't implement the measures that the measures are being used in programs that, you know, they're not aware of or they're just out for use, it's impossible to keep track of.

This kind of leads us into our discussion of, again, what are the expectations for this sub criteria? You know, the questions we're asking - the correct ones, are we - is there additional guidance that we should be providing or developers should be providing? Is there again, is there an option for justification or a plan to solicit and collect feedback for this measure? So I'll turn it over to Eddie and Rose to facilitate this discussion.

Rose Baez: Thanks Madison. I think (Sarah), you have your hand raised.

(Sarah): Yeah, thanks. So actually, in this case, I sort of feel like there might be - the expectations might be pretty different for the first-time submission versus sort of repeated submissions.

In the sense that from my perspective at least as a measure developer, you know, when you are developing a measure you really do need to make sure that you are addressing some of these issues around making sure that these are meaningful for end users and that you're getting - you're really are getting the feedback and incorporating it in your measure development.

I think so - like we've been saying you know, for resubmissions where as you pointed out, quite aptly, that the measure developer is often not the measure implementer, I think it becomes a lot harder to be specific, even if we try to elicit sort of feedback, you know, more generally. I think that's often - it's often hard to get.

Eddie Machado: I think Lee, I think your hand is raised.

Lee Fleisher: Yes. So I think that there's a little bit of how I think of guidelines, which I'm involved in writing and just as we pushed out, making sure we answered everybody's comments, so the question is and I know back to (Elisa) or NQF

staff in that as you get a more robust measure response system, or as you - particularly the comments from the initial reviews are out and we get those responses back from the developer.

A lot of times, and we've talked about this, closing the loop on the concerns from the previous review and what we hope to see in the next review, making sure that we'll look at for the comments from the committee, but also if anything comes in through the Web site. I don't know how frequently you get comments in. If in the future it may be more robust and what the developers ask and how they respond to those comments.

Ashlie Wilbon: So, this is Ashlie from NQF. I think there may be a couple of answers to your question. I think one of the questions was around what kind or the kind of volume of feedback we get through other tools that we have through NQF to get the developers. And we're going to talk about that in later sections, but I think essentially the answer is we don't get very much.

And there may be some, you know, various reasons for that. Maybe we haven't, you know, well publicized that we have a feedback tool. We do have commenting that's open during various points in the (CEP) process, but often times when we solicit comments on feedback or implementation, those are all also often minimal.

So I think that kind of closing the loop process is often - there's not really a whole lot to pass on. And I think the other part - maybe the other answers to your question if I understood it correctly, is that around what the developers are - the expectation for the developers? Do you have considered those comments and what they've actually done to adjudicate them? Is, that the other part of your question from, if I'm understanding correctly, Lee?

Lee Fleisher: Yes, very much so. Particularly the, you know, we try to bring back the comments from the previous review where we said, okay, we'll accept the feedback. This looks - accepted developers answer because it's the first of the initial submission for the first maintenance.

But how well do we track that and do we see whether or not it responded even by soliciting comments from the community?

Ashlie Wilbon: Right. I think that there's also, I mean, maybe there's probably some degree of improvement that we could do, kind of carrying over information from evaluation cycle to evaluation cycle.

And so there's obviously internal tracking systems that we have that we're always working to refine and make sure that information can carry over, you know, a six-year, nine-year period between, you know, over a two or three-year maintenance cycles.

So I think that's something that we certainly are working on. But I think to the degree that developers have the opportunity to address how they've provided either responses to comments that they've received or that they take an action on feedback that they've received, there is an opportunity in this submission form to describe how they have adjudicated the feedback that they have received.

Lee Fleisher: Appreciate it. Thank you. Because I do think you have opportunity.

Man: So I guess the second question, and maybe we weren't, he alluded to this about what guidance might there be for developers or additional guidance might there be for developers for collecting and submitting the data on feedback. Any other thoughts on that?

Melody Danko Holsomback: Well, I think it's - this is Melody. I think there needs to be a system of how to get it out to your national organization. Like until I was on the committee, I never received even in, you know, whether it's CMS or other news feeds, information about measures being reviewed or submitted. So like how does that information even get out?

And it comes back to maybe the advertising of knowing what's under review and where to comment unless you're specifically going in monthly focusing on NQF and that's kind of your job.

So looking at a way to get notifications or maybe a monthly calendar out to your bigger organizations like AMA and (NACO)'s and others that they can then, you know, with a link of where to comment that we can then distribute to our end users to bring it, you know, in front of them to know, yes, this is a measure being reviewed. It's very important that I've found burdensome or that I think is good. That I want to comment. That's a good, you know, either way. That we can distribute - help distribute that because I think if you're not meeting key people or key processes to get that message out. It's just going to sit within that internal hub of the developers and those that are using it for their, you know, to collect their quality measures and not get to the real point of who should be understanding it and understand the use of it.

Eddie Machado: Good point. Heather, I think you're raising your hand.

Heather: Yes. So, you know, I mean, I definitely think that measure developers need to take a central role in having some accountability for measure feedback. I would echo the comments of the other person that just spoke.

I think that a lot of providers don't know how or where to give those, but measure developers should, I think, be actively thinking about you know, whether there are medical societies or professional associations that represent providers that would be reporting those measures, you know, in many instances, you know, some measures are reviewed on an annual ongoing basis just for maintenance.

They could be tapping into members that sit on their technical expert panels to reach back out to providers or to at least get some message out to providers to share feedback about measures.

I think, you know, given the current climate and what's going on with, you know, just the quality of reporting programs in general. I think providers are really overwhelmed, but I think most of them would welcome an opportunity to give feedback on measures if they knew when and how to do it. And we could make it simple.

So my suggestion would be is if we, you know, give some basic suggestions about how that could be carried out prior to having these measure developers come forward, either for the initial review -- where I do think they probably have gotten a little bit more feedback -- but also just as importantly coming back for the, you know, maintenance and, and endorsement of these measures.

Really being thoughtful about doing a track - an environmental track to make sure that the measure is really doing what it should be doing. Whether that there are issues with implementation.

Eddie Machado: Okay.

Connie Anderson: This is Connie again that was eloquently spoken and I would wholeheartedly agree with that. I think you've really hit the nail on the head with the developers in the feedback. And the process to give the feedback.

Many of the providers and others don't know how to get the feedback or might not even know the measures are under development. So I think the onus of responsibility is to figure out how we can get the message results to people who can then provide (unintelligible) feedback the developers.

Rose Baez: I agree.

Jean-Luc Tilly: I almost wonder if one of the things would be to actually figure out a more standardized mechanism to - for developers to get feedback. Whether it be, you know, whatever the venue they put their measure in.

I, you know, I know its extra burden possibly to the users, but at the time of reporting, whenever they're doing reporting, whether it be they're doing it in batch or submission for like say, you know, quality payment and the annual base - submission basis. That's your opportunity to then automate that feedback mechanism back to the measure developer, because, you know, like many have said, most providers have no idea how to get feedback to measure developers on the current measures they're using.

Rose Baez: This is Rose for those of you that are measure developers do you feel that you have a good sense of those that are using your measures? I see (Elizabeth) your comment here about beta testing with end users, but I wonder you know is there is a gap in itself in measure developers in having that awareness of who was using those measures to be able to do that reach out?

Heather: This is Heather again. I think that really depends on who the measured developer is, where the measure potentially is being tested or used. So I can just speak briefly on our experience the American Physical Therapy Association has a QCDR.

So for us getting feedback on measures during development is really simple, and obviously if our measures get utilized by other providers that are using other QCDR registries, we can do outreach to those registries as well and try to get feedback from their users.

I think it gets a little bit more complicated probably for measured developers that don't or maybe are not attached a registry per se and I don't know if there is anybody on the phone that would have comments about that from that viewpoint.

(Sarah): This is (Sarah). Yes, I think it can be more difficult certainly. I'd say you also have an issue in terms of the measured developers are (ARC) or CMS. They are obviously big organizations, they have a much better ability to get that feedback. For smaller measure developers it's much harder to get that feedback even if you have a sense of the general population who might be using the measure.

Eddie Machado: (Trish) I think we have - you're raising your hand.

(Trish): Oh yes. Thanks. So just wanted to offer up you know a little bit to the conversation. As a measured developer we also implement and request data on the measures. So specifically as soon as they commission a part of our accreditation certification then on occasion their measures do get used within the CMS program.

Beyond that, it's sometimes challenging to know if other vendors have the registry or a quality measurement vendor has embedded measures into their software. Typically, we know because it becomes ultimately becomes part of our accreditation process, but we don't always know that.

So I think that's a challenge for measured developers in knowing how (light) spread the measures became - how well adapted in the field. We're able to track what we intend the measures for, but beyond that if others adopt them, it's just hard to get a complete understanding of where the measures may reside.

Eddie Machado: All right. Any other comments? Otherwise I guess we can turn it back to the (Youngcatt) staff.

Ashlie Wilbon: Thanks Eddie. This is Ashlie. I'm going to just walk through the first component of the usability criterion which is improvement. This sub-criterion really focuses on how the committee understands how much progress towards achieving high-quality care has been done based on implementing the measure.

Some of the key questions that we're asking the committee to consider as they define the criteria is whether or not specifically for an initial submission, if there is a credible rationale for improvement. So we're asking them to describe how they would envision improvements to be accomplished through using the measure.

And if the measure is already in use, keeping in mind that all measures that - even a new measure submission may already be in use. So often times if we get measures from (Thoro) partners or other developers who have been using their measure and the measure is being used in their other programs but it has

just not been submitted after endorsement before. So the measure even as an initial submission, may be still be a measure that's already in use.

So for those measures in particular, as well as maintenance measures, we're looking for them to describe how improvement in the measure has been demonstrated either through performance trends or the number of people they have identified and have received high-quality care as a result of the implementation of the measures.

And if they have not been able to demonstrate either of those things we're asking for them to - we're asking for the committee to considered what are the reasons it hasn't been done, has the developer been able to provide a credible rationale for why it has not been done or why there hasn't been in improvements in quality based on using the measure.

In the submission form, particularly what we're asking the developers is again for them to provide information that demonstrates that there has been progress towards achieving high-quality care and then again if not in use at the time of endorsement, then how would they intend the measure would be used to improve quality for patients.

And so again, questions similar to the theme of our other questions about how developers should be approaching, gathering this specific information around performance. I think we've had some comments already around this aspect, but again in the context of specifically in terms of improvement and demonstrating that the measure is helping to drive improvements in quality and what other stakeholders might we be able to reach out to, to target this information.

So I'll hand it over to Eddie and Rose at this point to see if there is any additional thoughts on this.

Rose Baez: This is Rose. I don't see any hands raised, but why don't I just open it up to see if there's any questions or discussion around this.

Ashlie Wilbon: Hi, I think we already kind of talked about this previously on the approaches whether it's you know reaching out through organizations or broadcasting or you know something routine. If there's a calendar like at the beginning we had the - there's a deadline for submissions and comments or renewing measures, getting those calendars with an easy to access link out to providers in various ways that they know what's coming up and how to address them, maybe that's through EMR vendors and since everyone who needs to be on an EMR, looking to the vendors and national organizations to push those messages out.

Eddie Machado: I'm surprised using American Heart for hospital stays - American Hospital Association, have you tried to push on people at AHA, (Nancy) and others to get comments to help or just take their comments? That's what I was...

Ashlie Wilbon: Was that for (NQSS)?

Eddie Machado: Yes.

Ashlie Wilbon: So this is Ashlie. I just wanted to make sure. So I think what we find is that there are folks within AHA who are tuned in to specific issues that are relevant to their members and so we do find that they engage particularly around hospital level measures that are going through the review process.

I think - the newer measures obviously we tend to have more involvement and then as things comes back for maintenance maybe not as much engagement because folks have been using it for some time, they're more familiar with it, unless there's been unintended consequences or other issues that have come up.

But I think for some of the more larger groups where there are measures that are relevant to their membership, we tend to get engagement through folks in the organization who are focused on a (QFR) for other channels that we provide.

Woman: I guess my question is, I think if you're giving the question more to the end user, you're going to get a little more specific rather than for example, if it's a process type measure, an EMR system can make it easy for the process, you're not going to get a lot of feedback on that. They're just kind of doing it.

But is it really making a difference and I think that's what we want to get to in the end is are these measures making a difference for you population or is it another checkbox that you have to do? Because sometimes if it's an easy checkbox, they're just doing it and you're not getting feedback.

Rose Baez: Right.

Eddie Machado: Are there other channels that folks believe that should be maybe looked at to try to gather information maybe more directly related to specific programs in which the measures are being used whether it be federal, state or private programs? I know there's a challenge in trying to get the folks who are operating them to work around this, but I'm just wondering whether that is also another channel that should be looked at because there's an instant context there for the folks who are working in there.

Woman: Yes. I agree with that - that's an important aspect to look at.

Eddie Machado: Okay if there's no more comments, maybe we can keep moving along.

Man: Great. Thanks. So then we're just on Slide 45 now if you're following along using the numbers. So when we're thinking then about usability and really this is all about unintended consequences. For a variety of reasons we rarely see anyone reporting unintended or unexpected benefits choosing a measure.

Typically it is using performance measure you get what you strove for. But there are occasionally unintended consequences or either the developers - generally the developers are anticipated ahead of putting forth the measure but sometimes they'll get back as feedback.

And so this criterion is really just trying to assess the balance of those. Do the benefits of the measure outweigh the harms that are in angst or observed. So the key questions in the evaluation of this criterion is that you'll find is really a validation criterion and the current document, first identifying those unintended consequences, identifying the unanticipated benefits.

As I said, it's really quite rare. And then deciding do those unintended negative consequences outweigh the benefits. And then in the submission form it's pretty straightforward really, we just ask the developers to report in two different fields. \

One, we just ask them to explain any unexpected findings, positive or negative, observes is part of the implementation of this measure, specifically as to its impact for patients and then explaining any unexpected benefit in sort of the same way.

So the discussion questions here which are of course we'll lead you through is when we're thinking about unintended consequences and other ways to address questions - who is the target audience for this information? How does affect how we're asking the question and other ways to identify unintended consequences such that they could be incorporated into the endorsement and evaluation. And with that I'll turn it over to Eddie and Rose.

Rose Baez: Thank you. I do see a question from (Matt) coming through the chat. Do you want to ask your questions?

(Matt Davis): Yes, this is (Matt Davis), I didn't know if I would be able to be heard on this call. So I was just wondering, I work with patients with spinal cord injuries and we've been advocating regarding the (Cardi) quality measure and have seen some unintended consequences as (Bole) removal and centers that are not aware of how to manage (unintelligible) and bladder in these patients.

And I'm just wondering when a very specific small subset of the population comes forward with a concern, is there ever a role for really kind of reevaluating risks and benefits just for that very small population? So like review of the validity of the measure of that population. Review of - are the benefits the same for this population as you would expect for the general population?

Ashlie Wilbon: Right. This is Ashlie from NQF, I think that's a real interesting question. I'm trying to think, I've been here a long time. I'm just trying to think back on whether or not that specific issue has come up. I can't think of any examples in my head - off the top of my head, but I'm considering, it would take us kind of looking at the measure population that was actually specified in the

measure to see whether or not - how the measure is defined that that population would be included.

But also we do have a process built-in for what we call ad hoc review which we haven't mentioned here but probably should have around it's an opportunity either for the developers to identify something that's changed in their measure that would warrant an off cycle review or a very focused review on the measure around unintended consequences.

Or a member of the public would come forward and say we've been using this measure and we've noticed that there's these unintended consequences and therefore we think that the measure should be reconsidered for endorsement or that the developer needs to make certain adjustments to the measure in order for it to be continued for it to continue to be endorsed.

So we have had those situations, but I think they've been broader around unintended consequences for the broader populations or around specific changes in evidence or guidelines that would kind of change the way the measure is directing practice in a certain way, but I think that's a good question. I can't think of any examples.

I don't know if there are others on the phone who have been involved with NQF that can think of a specific example, but I think it's a consideration.

(Matt Davis): Well this is a specific example. I mean because we've had presidents of national organizations who commented on this and we turned in a letter last year that was signed by 10 different presidents of national organizations that were requesting an ad hoc review.

Ashlie Wilbon: Oh, okay.

(Matt Davis): And so I'm just trying to figure out how do we flip a switch and how do we determine what's required to kind of flip the switch to say hey you know this really deserves a really in-depth review of all of the relevant data that's out there because that's pretty fuzzy. I can also give you some feedback on how to navigate your Web site or challenges that I've run into because we've been doing this for four years.

And maybe this is more than you can talk about right now, but yes this has been going on for a while.

Ashlie Wilbon: Okay. Yes. So I didn't understand that to be tough questions related to specific issues. So without knowing kind of the history of that, I'm reluctant probably to comment on exactly what the process that was employed to kind of adjudicate the ad hoc request.

So I can't speak on that, but we certainly I think for now, potentially if there are specific comments you have on navigating the Web site as it relates to submitting measures, I think we'd probably like to hold the committee's discussion on that for now, but certainly if there are other points where that fits in.

I think in the next segment we're going to be talking about the feedback tool and some of the other options or channels that we offer through the NQF Web site to submit feedback.

Madison Jung: Hi this is Madison with (Interest). I would just say that if you have any other additional comments (Matthew), feel free to email us at our email address or we can take more questions at the time of public comments.

(Matt Davis): Yes. I think that would be good. We may be able to touch base offline at another time.

Madison Jung: Great. Thank you.

(Matt Davis): Okay. Thanks.

Eddie Machado: I think we see (Connie's) hand is raised.

Connie Anderson: Sorry I was on mute. Yes in terms of the targeted audiences for soliciting information about the unintended consequences, I think it's a measures developers and it is who is implementing those measures and there are all types and I think somebody mentioned this earlier of societies and things that collectively could provide information about how these measures have been implemented and what those unintended consequences are whether it's for our organization and the nephology the American Society of Nephology, is a great example.

But there's a lot of others and I think it's very important that you go back to those that are implementing these measures to solicit these methods, how the measures have gone and what the unintended consequences have been if any or what the unintended benefits have been.

And I think that's important feedback to the measure developers to hopefully help maybe define measures and improve the measures so that you can achieve the high quality of care that you find to accomplish by the measures.

Eddie Machado: Okay. I did see we had (Elizabeth). I think you made a few comments in the chat box. The last one I think being related to the star rating system. Do you want to raise that to the group? I think it may be relevant again for the

purposes of unintended consequences. All right any other comments from the group on this topic? No? All right. So I guess we'll hand it back to Ashlie and the folks at NQF.

Ashlie Wilbon: Thanks Eddie. So the last couple of slides in this section are just to kind of summarize our discussion on some of the other specific use and usability criteria and submission items but also bring your attention back to some of the other sub-criteria within or criteria within our evaluation criteria that also solicit information to the developer related to feedback for the measure.

So with any important criteria within opportunity for improvement we're asking them to provide performance scores on the measure to demonstrate that there's still a drop in performance or an opportunity for those being measured to improve and then within feasibility we're asking information.

We're looking to evaluate information around the data collection strategy and the availability of data, missing data and how that has affected the implementation of the measure and considerations obviously for patients for measures that require the use of a tool that interfaces with patients.

So the last couple of slides again are just kind of overall discussion questions based on all of the criteria that we reviewed up to this point and really kind of looking back at these six sub-domains if you will that we are using within the criteria to solicit information from the developer on feedback they've received on their measure.

So within the use and usability criteria, we're asking about accountability and transparency improvement to benefit outweighing the harms of the measure, about feedback from those being measured and others. Within feasibility

we're asking about the data collection strategy and the importance criteria we're asking about the opportunity for improvement.

So just kind of again thinking about all of those questions that we've discussed, the different topic areas within those criteria, is there any other kind of domain or kind of type of feedback that we should be considering that would be useful to the committee in their evaluation for endorsement? And I'll hand it over to Eddie and Rose to talk about that.

Rose Baez: Melody I see your hand is raised. Did you want to start us off with a question?

Melody Danko Holsomback: Yes. For another category, the feedback and those things measured and others, well I guess it wouldn't go into that one, but when looking at the measure, what are the expected outcomes? I guess I don't know exactly how to put it, but to look more when you're developing the measure is it a process improvement versus an outcomes measure?

How can we get the measures to be more looking at what are the outcomes of implementing this measure and how to do that and measure it appropriately rather than just doing a measure to doing it? When they're developing these measures looking at more specifically the outcomes rather than the process of doing it and how to measure those.

Rose Baez: I think there were comments earlier about you know identifying the intended use of the measures. So maybe along those same lines about you know the outcomes related to that specific intended use of the measure.

Melody Danko Holsomback: Yes because I think a lot of - when we talk about the burden of measures and trying to collect the data on them, if you're looking at certain

measures and I'll take this out of the same things that we've experienced with our different pay-for-performance.

But if you develop measures that have the intent is to lower specific outcomes rather than to check each patient, was this task done or to check the sample of patients, are you looking at those to cohort the patients that it applies to and what their outcomes were as far as claim space for admissions, admissions or certain procedures to be done or not to be done, those type of things and how can the measures be more on that outcome portion rather than the collecting of the individual tasks that were done within them.

I think it puts the responsibility on the providers to do the right thing but without having to have them document that they're - specifically that they did this one right thing and it's the overall care of the patient.

Eddie Machado: That's a good point. Other ideas on additional domains or reaction to the existing domains? All right.

Rose Baez: I think the existing domains look great overall. It seems like they cover what we would want covered.

Eddie Machado: Okay. Great. All right we're two hours plus into the meeting. So I think though even though, I'm wondering whether also tiring out here on this topic. I don't know we'll throw it to the NQF staff and see where we want to go at this point.

Man: We had assumed that evaluation is also part of or it's one of the domains.

Eddie Machado: So your comment is about your suggesting that evaluation was one of the domains, but...

Man: It's a current evaluation domains include all of those headings, but I think the actual physical evaluation if you want quantitative evaluation or qualitative evaluation would be one of the domains in itself.

Eddie Machado: I see. Okay.

Heather: This is Heather. I don't know if this is included in the evaluation domains, but the overall impact of the measure and how the measure has trended if that data is available over time I think would be worth looking at. I'm sure that maybe fall under improvement, I don't know. But I just wanted to point that out.

Eddie Machado: Yes. That's a good point. I mean it's probably somewhere between improvement and the importance criteria I think upfront. Other comments from folks? Okay. I guess we'll hand it back to Ashlie and the rest of the NQF staff at this point.

Ashlie Wilbon: Thanks. I wonder we kind of do a pulse check. There's one more discussion slide we have in this section that's a bit of a kind of a summary discussion kind of the overall structure of the use and usability criteria and it not being a must pass. Do you guys feel like you have it in use to talk about that now?

We can certainly save it for our webinar next week but I guess maybe a pulse check to see how folks are feeling. If you have thoughts on that now, we could maybe discuss it now and just start fresh with the next section on the next webinar or if you'd like to kind of think about it and start here with the next webinar we can do that too.

Does anyone have any strong feeling either way? Co-chairs would you like to make an executive decision either way?

Eddie Machado: Oh, I think Mark has his hand raised.

Ashlie Wilbon: Oh, okay,

Mark Huang: So I guess. I'm just going to do a quick comment on just usability much like we were talking about in the other piece about use and as a must pass - it is, it's a must-pass but there that three year, six year, sort of landmark and then you talked about that one slide where it said you know what about exceptions of explanations - this might be another one of those that maybe you know you could include that. Meaning like if somebody - I almost feel like, maybe we want to think about as a must pass, but then give them an ability to explain it if it's not meeting whatever criteria it's come up with.

You know much like you didn't have unusual circumstances regarding use, you could have potentially unusual circumstances around usability. Maybe you haven't seen as much improvement in the measure as you would have anticipated, but maybe you need to revise it some way or fashion. Does that make sense?

\

I mean I'm just trying to think about - I think usability is important in the context of that, but maybe it's just a matter of we need to give it an out if it's not really panning out as initially anticipated or maybe it's like if you have unintended consequences maybe you want to expand that further to either limit or prevent those unintended consequences in future measure modifications.

Eddie Machado: Thoughts to consider. I mean I think it's not black and white. I think that's a better way to think about this. Others? Additional comments? (Jan) I see you have raised your hand.

(Jan): Yes. Again, I'm thinking in terms of the valuation framework that all the domain at the top has been listed but we can come up with a sort of flow if you want or the natural steps in this evaluation process. For example, the word before last, the data collection strategy, which is very strong, I think based on that there should be one other step maybe or maybe it's included already in the other data analysis and because I think that - that data analysis would inform the other steps in terms of improvement or opportunity for improvement.

And then if it's an - what I'm trying to say is or suggest it is to sort of come up within a valuation framework and think of other things that can be added to that. If you think it should be part of our role.

Eddie Machado: I don't know. (Steph) could you comment on that? I mean do you see that as part of the group's potential role or is that a little bit out of scope on this?

Ashlie Wilbon: Hi, this is Ashlie. I just want to make sure I understand the question. Is it about changing the order in which they're asked throughout the evaluation criteria? I'm not sure I quite follow the question. I mean I heard a couple of things in there. Can you just clarify for me?

(Jan): Sure. It's not exactly the order, but it's just to sort of create a flow if you want of the steps to create an evaluation framework and so what I was sort of highlighting was that maybe another step would be data analysis and then how this can inform all the other steps.

Ashlie Wilbon: Okay. So I think it sounds like, I probably wouldn't describe it as an evaluation framework, but maybe it would maybe guidance from the developers in how to approach kind of providing a story about their measure and how the feedback was collected and how it was used to improve the measure.

But I think certainly that sort of recommendation would be welcome and I think we can try to frame it in such a way so that it is I'd be reluctant to change the criteria like kind of wholesale in that way, but maybe to provide some additional guidance around some of the existing domains or how developers should be thinking about it.

And maybe including other steps or other domains that helps them kind of build the case on how they've used feedback and try to build the case of the usability of their measure. Does that sound in alignment with what you are describing or am I totally off?

(Jan): Yes. You are exactly on it.

Ashlie Wilbon: Okay. Okay.

(Jan): Thank you.

Rose Baez: Ashlie this is Rose. I'm wondering does that fall more within that improvement domain where we're looking at performance trends and being able to tell the story of how the measure demonstrates improvements.

Ashlie Wilbon: I think it certainly could. I actually wanted to ask I think (Jan) who was describing kind of the empirical justification for the measure. Can you talk a little bit more about that?

Is it something beyond - it sounded like what you're describing was something beyond the kind of improvement results, but kind of to demonstrate empirically that the improvements, that implementing the measure actually was associated with the improvements in quality healthcare? I wasn't sure if I was quite making the connection about this like additional step you were suggesting.

(Jan): Yes, no.

Heather Smith: This is Heather.

(Jan): Sorry go ahead.

Heather Smith: Sorry. This is Heather. I think that might have been me that spoke up about that. I think it would be really helpful to look to see how the measure has performed over time to try to discern what the impact of that measure was on the target paradox. That's exactly where I was going with that.

Ashlie Wilbon: Yes. So that sounds like a little more than I think that we currently ask for, but I would certainly be interested in kind of hearing more either on a subsequent call or folks have the appetite now to explain just so - the staff make sure we understand it so we can explain it. But kind of what that would look like using what we already collect.

So for improvement we're collection information on trends. So like show us how over time this measure has performed so we can see where the gaps in performance are so between you know 2013 and 2014 or in 2015 for example over three or four years of data. Has the gap narrowed? Is there still a significant gap?

You know I think we also ask the series in population stuff like that. And then in the improvement section within usability, we ask them to talk about the improvements that they believe have been gleaned from implementing the measure. So I guess I'm - if you could just maybe tie those together for me, just connect the dots for me just so I can understand how that would fit in I think that would be helpful. Would it be something different than we're already collecting?

Heather Smith: No. I think that you are collecting what you should be. I was actually asking if that was part of the process. It wasn't clear to me in looking at the categories originally if that was included. But I mean I would like the measure developer to be thinking about whether those trends - the improvements that they're seeing are in part due to the measure.

I mean that's the goal right? The goal is implement the measure to improve the quality of care and I guess on a counterpoint to that what happens if that measure just stagnates? So it's stays at a certain level and doesn't improve over a number of year, what then happens to that, even though there's still room for improvement, I guess?

And I guess my thinking along the lines is - let me just tell you where my brain is right now is I think we've got a lot of measure development going on and I guess you know I would question whether or not there are better measures in the field or the measure needs to be tweaked so that it can be more usable.

I guess I'm just thinking for a measure that's kind of an average performer is it not quite getting to the quality improvement that we want. Are there other issues that are holding providers from being able to get those improvements?

Is it just the process that we have in healthcare? You know I would just be curious how we handle kind of those average performers?

Ashlie Wilbon: This is Ashlie. I think it was a great thought and I'm wondering if you think it would be different for outcome versus like a process or an intermediate outcome measure where you know trends over time would be showing actually the outcome versus like a process measure which may be more approximate against the trend - looks like the outcome that we want to see and whether or not we need to be more explicit about that connection within that section.

Because I'm not sure that we are. So I'm just thinking out loud, but I think that would be helpful.

Heather Smith: Yes, I think it would be again, part of the purpose of having the committee is to really think about how we can improve feedback to make these measures more valuable and really function at the highest level. And so I would challenge us to think about, you know and it is difficult.

Going back to earlier comments about you know having measures available for providers, enough measures to be able to navigate some of these quality programs, quality reporting programs like something like mix now, you have to kind of finely balance that with I think also trying to evolve the science of measured development and really trying to get the best measures out there that are going to do something.

Because I don't think as providers or as patients, you know we want to have measures that are meaningless or are not really hitting the mark. So I guess I'm kind of putting that out there.

I'm not really sure how to operationalize that, but I just think we should be thinking you know is this measure really helping us to get to where we want to go, and if not, how can we make that better and has somebody else done that in a measure that's not NQF endorsed?

And I realize that's probably a bigger ask, but I think we have to start asking those questions as we're starting to see more and more measures get developed potentially by QCR's and maybe not all of those measures are coming forward right now through NQF while also looking at measures that have been around for quite some time and trying to figure out that balance.

Ashlie Wilbon: That point is well taken. Thank you.

Eddie Machado: All right. Any more comments or thoughts on the questions listed on the discussions line? All right. I don't know Rose, I'm thinking maybe it might be a good point to sort of stop and wrap-up in anticipation of the meeting next week, maybe?

Rose Baez: I agree. I'll just thank everybody for their participation and for the lively discussion and I look forward to more of it next week.

Eddie Machado: Ashlie or others, do you guys have any parting words or anything we should be aware of?

(Addison): This is (Addison). Again I would just echo the thank you. I think - I agree we are in a good place to stop. So I won't do the full recap that I usually do. I'll include that in our slides for next week, but just to recap at a very high level.

So what we've done is kind of gone through right now the use and usability components of our CDP process. During next week's meeting some of the

topics we'll cover is just discussing some of the other feedback that we get through (Entrust), normal processes we touched upon it a little bit in the environmental scan, but just highlight what goes into the committee's evaluation.

And then I think the bulk of our meeting next week will be used to describe the challenges and gaps in what we have and kind of start thinking about our next steps moving forward.

So unless there are any other questions from the committee, we can turn it over to (Matthew) for comment. So I'll pause. Are there any questions from the committee just about what we talked about today?

Okay. Hearing none and as always feel free to email. I'll turn it over to (Matthew) for public comment.

(Matthew): Thank you Madison. So at this time, we will unmute all the lines in the web meeting. So if you would like to make a public comments please press yourselves or if you would not like to make a public comment, please place yourself back on mute.

Operator: The conference has been unmuted.

(Matthew): All right. It looks like we do not have any public comments for the committee. Just want to reiterate what everyone said thank you all for joining us today. So next steps we will be rejoining for a web meeting number 4. This will be next week on Tuesday, May 7 from 2:00 p.m. to 5:00 p.m. Eastern Standard Time. Thank you all.

Ashlie Wilbon: Thank you everyone.

Woman: Eddie or Rose anything to add before we head into next week?

Eddie Machado: No, just thank you for your time and we'll talk next week.

Woman: Okay. Great. Thank you. And Eddie and Rose if you could join the subconference, for a quick debrief, that would be great.

Eddie Machado: Okay.

Rose Baez: Okay.

Woman: All right. Thank you everyone.

(Jan): Thank you.

Operator: Self-conferencing is now active.

END