NATIONAL QUALITY FORUM

Moderator: Sheila Crawford
August 16, 2012
2:00 p.m. ET

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

Reva Winkler: Good afternoon, everybody. This is Reva, Alexis and Adeela (at) NQF. Thank you all for joining us today for this conference call of the Infectious Disease Workgroup – Workgroup B.

And so today on our call, the workgroup members are Kathleen Brady, Sue Elam, Thomas Giordano, Jeffrey Beal, Aaron Milstone, are any other steering committee members joining us today as well? OK.

Some of the other members of the steering committee, anybody is welcome to join us. And so from HRSA, the measure developer, we have Marlene Matosky and from NCQA we have (Donna) and (Jenna). And so I think we can go ahead and get started.

Thanks, everybody for the workgroup for doing your preliminary reviews. We sent you yesterday the summary documents that really described the input and summarizes your preliminary evaluations on the criteria for the six measures that are assigned to this workgroup.

So, we will be showing that document on the webinar and we will also have it in written form. So, that's going to be one of our main resources for the discussion today and what I'd like to do, we did this on the first workgroup call yesterday and it worked out really well is for – we'll go through one –

each measure at a time and then, as we talk about each measure, we'll go through the criteria in order and see how the workgroup thinks, whether there are disagreements, whether there are questions for the developers, whether there are any issues that you want to raise about how well these measures meet the criteria.

So, are there any questions from anybody about the documents we're going to be using and/or what we're intending to do? OK. Super. So, the first measure on our agenda is Measure 404 which is HIV/AIDS CD4 cell count or percentage performed, and this is the percentage of patients age 6 months and older with the diagnosis of HIV/AIDS with a CD4 cell count and percentage performed at least once every six months.

So, Sue Elam, I think this is your – oh, I'm sorry. Kathleen Brady, this is your measure and would you like to kind of start us off, summarizing your thoughts and the input for the rest of your colleagues on impact and performance gap?

Kathleen Brady: OK. So, based on – as you can see from the summary that for impact, all members determined this indicator to have a high impact and there were some comments regarding performance gap. For that one, the voting was two for high and three for moderate.

The comment was this measure was used in the 2009 and 2010 CMS Physician Quality Reporting System – or actually, this is the information actually that was reported by the developer. For this measure, the average performance rate for eligible professional was 76.8 percent in 2009 and 83.9 percent in 2010, and the developer felt that those numbers indicate there was a gap in care with room for improvement, and it was noted that the measure's not stratified by patient groups or cohorts that could potentially be affected by disparities in care.

And the comments for this group that there was less than optimal performance with the statistics noted indicating that there really is improvement, possible there is a gap and there were two comments from reviewers regarding the lack of disparity information.

In terms of evidence, based on the decision logic, four members agreed that this met the criteria and one did not. In terms of the evidence, four reviewers thought the quantity of evidence was high, one felt it was moderate, one was – the comment was that for the person that gave it a moderate vote was that the studies were not randomized controlled trials. The body of evidence was greater than five studies with large numbers of patients.

In terms of quality, there were two members who gave the – reviewed it as high quality, three is moderate. There weren't any specific comments provided for that. And in terms of consistency, three members voted that it met high – the high criteria for consistency, one for moderate and one felt it was insufficient, and there was really no specific comment for the reviewer who found it insufficient.

Reva Winkler: OK. Perhaps, whoever thought that might share your thoughts with the rest of the group?

Thomas Giordano: This is Tom. I marked it insufficient because I didn't think that there was quantified estimated benefit which I thought was one of the instructions. And I think this could apply to a lot of the measures. You know, clinically clearly, it's beneficial to have your CD4 monitored but with the data presented that people who don't have it monitored do worse. And that's why I said it's insufficient.

Reva Winkler: OK. I mean, the criteria do focus in on, you know, the quantity of studies on the relationship between the process of care which is measuring CD4 counts and patient outcome. And so, certainly, there's a lot more information around the body of evidence when you look at net benefit. But the criteria actually do focus around the quantity, quality, and consistency of those studies.

So, you know, certainly, we'll leave it in your hands to assess that issue you feel most appropriate.

Thoughts from anybody else from the committee? So, in general, are there major concerns about whether this measure meets the importance criteria, from anyone on the workgroup or the committee? OK.

In the absence of comments, Kathleen, why don't you summarize your thoughts on scientific acceptability, reliability, and validity?

Kathleen Brady: OK. So, overall, in terms of scientific acceptability of the measure properties, all of the reviewers voted yes on that and it met the criteria. In terms of reliability, there was one person who evaluated it as high as four as moderate. And the comments were that the measure is well defined, implementation is straightforward although it was unclear how this specific measure performs in other health records.

And so, for validity – and actually, I think since the specific individual components of the measure were not evaluated only the measure as a whole, I think that's why most of us probably gave it a moderate on reliability.

In terms of validity, all five gave the measure a moderate score and there was a comment about validity testing and face validity discrepancies but I'm not sure what the individual meant by that.

In terms of usability, four votes …

Reva Winkler: Can we just stop right now and let's see if anybody else in the workgroup has anything to say about reliability and validity?

Kathleen Brady: Sure.

Aaron Milstone: This is Aaron Milstone. I had a question, I mean, and this applied to a couple of these measures. I mean, for me, it doesn't like look this was developed initially within a small group of medical centers within one geographic area that all shared the same electronic health records. So, when you talk about reliability and validity, I was – I mean, I understand that if you're using an EHR reliability you think would be good if you're pulling it from the same EHR.

But they didn't look in multiple electronic medical systems. They didn't look in multiple geographic areas. Now, I know that CD4 count is consistent and it's not usually quoted differently but I was just wondering if others had questions about how that – or if someone can give me some more insight as to

the standard for measures that are only validated in small areas that hasn't been looked at more broadly, and whether that raised any concerns?

Kathleen Brady: I actually – it's Kathleen. I had a question about why that is because this is actually a measure that's been adopted by HRSA. So, I feel like there should be more information available than what was submitted by the developer.

Aaron Milstone: Yes, and personally since I don't know any HIV carrier, I didn't know that. So, I think that's available but it wasn't known to me that's why I was just questioning the – anyway, if other people have thoughts or more experience with that.

Reva Winkler: Do somebody from NCQA want to comment?

(Jenna): Hi this is (Jenna) from NCQA. You certainly do bring up a good point about the small groups of sites with the same EHR that is just how we were able to develop our test plan at that time. We don't actually have anyone from the testing organization on the call. These measures are developed with – jointly developed with AMA-PCPI and they lead our testing. We do plan to have them available to answer testing questions during the steering committee meeting, and we certainly can ask them any follow-up questions about testing outlines as well.

As far as the measure is being used by HRSA, in my understanding and we do have representatives from HRSA on the phone so they can certainly jump in, is that they do have measures that align very closely with ours. They are not technically the ones the NSQA owns. So, we can work with HRSA offline to see if we can get access to that data but, you know, that data does belong to HRSA and it is technically a HRSA-owned measure, not an NSQA-owned measure even though they might be very similar.

Reva Winkler: Any other comments or thoughts on reliability and validity?

Kathleen Brady: I had – I actually – I had made a note about this, about the comments. The difference between scores, between the automated collection of performance and the manual collection likely resulted from some confusion about the

numerator inclusion criteria, which codes should be used and the timing of the CD4 counts. And so I wanted a little bit of further explanation regarding that.

(Jenna): That's a good question. So, the comment about the coding was that there was actually, at the – when the measure was tested, there was this CD4/CD8 ratio code that was included in the list of codes. We since had removed that because the CD8 ratio is not an appropriate CD4 test to do. So, we have addressed that particular concern.

The timing comment was about – because we say within six months whether that meant within each six-month's period of the year or if it meant every six months. We did not change the language of measure to address that but would be open to making that more clear.

Reva Winkler: OK. Any other comments or questions on the liability and validity from the workgroup? Then, for NCQA, I have one question that was prompted from a little bit of the discussion we had on some of the other measures yesterday. When you are looking of the CD4 cell count performed, are you actually looking for the laboratory result or note that says it was done or are you actually looking for the lab test result?

(Jenna): For this particular measure, because it is specified for claims and administrative data and CPT 2 codes, it would be just a claim; either a claim for the CD4 test or it could be a note in the medical record. We don't require that there has to be a result segment.

Reva Winkler: So it could be an order that a person – it was ordered for a patient but they never went to the lab, and that would meet the criteria.

(Jenna): It could be. Certainly, if this measure was ever picked up and developed as an eMeasure in the future that gives us more opportunity to look for results. But I think, again because it's a claims and CPT 2 code measure, we haven't required the results to be present.

Jeffrey Beal: Well, this is Jeff. I would certainly hope that that changes and that they are required to prove that the result is present.

(Jenna): That's certainly a good point we can – I'm noting that now and we can discuss that internally.

Sue Elam: And I would agree – this is Sue. If you were going to be using this as a quality tool, we definitely need to have results, not just that the test was ordered.

(Jenna): I agree.

Aaron Milstone: This is Aaron again. Just so I understand the process more clearly. So, given those comments, does that mean that it changes the way we would view the measure as is?

(Jenna): Certainly it might, and that's really what – the kind of input we need from you all as given those caveats, the reliability and validity of the measure.

Thomas Giordano: This is Tom. If this is a process measure, then the initial process is complete, right? The provider ordered the test and acting – getting the result and acting on the result are important clinically. But are they – are they what we're trying to measure here? Verifying that the result is in the record is a lot more difficult unless you've got a data – a live data system that can – that can talk to your EHR which most – I think most electronic medical records have but it does change the nature of the measure.

Kathleen Brady: Tom, the only thing I would say about that – it's Kathleen, is if you have a patient who you send off labs on – you know, you send them off to the lab and they don't get them, I mean, do you just say, "OK, no problem. I'll see him next visit," or do you make some sort of contact to find out, you know, whether your labs – can we make sure you get them done, you know?

You know, there has to be a process that when you order something, you're looking for the results and if you don't get one, there's still an action.

Thomas Giordano: Yes, I agree. I mean, I think that makes sense. At least what we use are future orders and it's a lot more difficult to track a future order. You know, you have someone coming back in six months, and miss their lab visit; you

don't start tracking down until they miss their pre-clinic lab visit – you don't start tracking down until they miss their appointment.

In which part, I guess you would – you would intervene to try to get them into the labs and in another appointment, so. Sure, it's a higher standard.

Reva Winkler: This is Reva again. The reason I pointed it out is I noticed the title says performed in the description. And so, we need – you know, it's important that we're conveying the proper information about the measure.

Kathleen Brady: So, with that said, if it says in the description, performed, then wouldn't you want the data to support that it's the result that people are looking at and not just whether or not the lab was ordered?

Reva Winkler: Any comment from NCQA?

(Jenna): I didn't realize that question was directed to us. Like I said, we're definitely willing to take this question back to our expert panel that probably would not happen before this steering committee but we certainly appreciate these comments. It is a good point that the title says "performed" and that we haven't specified results.

So, like I said, we can discuss that with our expert panel and see what they think about – actually, looking for documentation of results.

Jeffrey Beal: And this is Jeff. I just have one additional thought which is this measure kind of bleeds over into the direction that I see HRSA going in trying to define that someone is retained in care which also strengthens the need to know that that CD4 cell count was actually obtained.

Reva Winkler: All right.

Thomas Giordano: This is Tom. Before we leave that topic completely, I think the issue of the ambiguous numerator definition is important to address. You know, you could – depending on how you interpret the numerator definition, you could the standard – you could this in two ways. You could operationalize this in a number of different ways.

One is, the patient have a CD4 count, in this case ordered, let's say and then, you simply look and see, do they have another one in the same time period – I'm sorry, in six months, at least once every six months. But is there some defined gap that's supposed to happen between those two visits or if someone has a CD4 done on day 170 of the measurement year and then another one done day 190 of the measurement year, does that meet the standard or is there something that has to – some space that has to – some time that has to pass between those two measures?

Kathleen Brady: Well, the denominator is people who had at least two medical visits during the measurement year with at least 90 days between each visit.

Thomas Giordano: Right. So, those are the people who are retained in care by at least some definition.

Kathleen Brady: Yes.

Thomas Giordano: But the numerator, you could have two CD4s performed relatively close together and meet the standard if you define it one way. On the other hand, if you say, well, I'm going to do it at least once every six months, meaning they get it once and then you have to wait six months before you count it again.

I guess, I just – I see a lot of ambiguity in the numerator definition that I think should be clarified.

Jeffrey Beal: I agree. This is Jeff. Why can't it read just like the denominator saying the CD4 cell count of percentage that's performed at least 90 days apart within that period of time? They could mirror each other and it really wouldn't change it that much. And I'd be more comfortable with at least 90 days between two CD4s.

Thomas Giordano: And this was cites as the reason or at least one reason why there was difference in the chart review data compared to the electronic data that people were misinterpreting or variably interpreting the numerator.

Kathleen Brady: Yes. I think based on that, there definitely needs to be more information provided about the definition of the numerator.

(Jenna): This is (Jenna) again. I hear these comments. I think they're important comments and I would say speaking right now again without referring this to our expert panel, that the intent would be every actually six months in between because that's more – that's more aligned with the guidelines. But like I said, I'm happy to take that back to our expert panel and work on some clarifying language.

Kathleen Brady: I just looked up the HRSA manager is actually at least 90 days apart. So, I would favor having it be more consistent with the HRSA measure.

Reva Winkler: All right. OK, any other comments for the reliability and validity or can we move on to usability and feasibility? Kathleen, your thoughts?

Kathleen Brady: Oh, OK. Sorry, I had you on mute. For usability, four reviewers ranked it as high, one moderate. The only comment I think was is that CMS is currently using it and it was used in the previous three years. There's been 2.7 percent improvement and so it is being used currently.

And for feasibility, it ranks high for one person and moderate for four. And the major comment there was regarding the coding errors. And I think some of the things that – the confusion that we've talked about was actually I think I may have been one of the people that voted moderate on that and that might be even my comments with the issue about the coding scores – coding errors.

Reva Winkler: Thoughts from other workgroup or committee members? OK. Any other comments on this measure before we move on to the next one? All right, well, the next one on our list is Measure 406.

Male: When you do the final, do you review the – do we rediscuss the preliminary assessment as suitability for endorsement given the discussion or?

Reva Winkler: Yes, and remember these were just your preliminary for you all to kind of share and talk with one another. What we're going to be doing is summarizing all of this that at the in-person meeting, we're going to go

through this again with the entire committee and you will vote formally on each of these criteria after having the discussions.

So, this is just a chance for you all to kind of share your initial thoughts. And so, how you voted now is not binding in any way. And if you want to go back and change some of your comments or add in to your comments, we'll be happy to help you do that so that the summary we provide for this steering committee meeting reflects what you want to say.

Female: But the comments that we've discussed here will be included in that?

Reva Winkler: Yes, do you see at the bottom of this document where it says, Workgroup Call Summary?

Female: Yes.

Reva Winkler: We're going to start filling that in.

Female: Perfect.

Reva Winkler: OK?

Female: Yes.

Male: So, it sounded there'll be a strong message that's there's concern about the measure after the discussion as worded in the numerator.

Reva Winkler: Yes.

Male: OK. Thank you.

Reva Winkler: Also, we will have both the transcript and a recording of this call, and we're able to use your exact language in writing the summary. So, that's our goal, is to really, you know, very explicitly reflect the discussion and concerns that you've had here and kind of help you remember that when you go and make your presentation at the in-person meeting.

OK, anything else? All right. The next measure is Measure 406. This is adolescent and adult patients who are prescribed potent antiretroviral therapy. This applies to patients 13 years and older with a CD4 count, I think is less than 500.

Female: But just so you know that what we received actually says, equal to 500.

Reva Winkler: I know. And that what I think that might be the way our program works. And that's one of the things I wanted to verify with NCQA that it should be less than 500 and that's just the translational thing in our program, and I'm sorry about that.

And then, otherwise age 13 and older, we have a history of an AIDS-defining illness regardless of CD4 count or who are pregnant regardless of CD4 count. And then of those patients, how many were prescribed potent antiretroviral therapy.

Female: Sorry, I just wanted to clarify it. It's meant to be less than or equal to 500.

Reva Winkler: Right. OK.

Female: Thanks. Thank you for that …

Reva Winkler: OK. We may need to do some editing to clean some of those things up. So, we'll get. All right, Kathleen, this is your measure again.

Kathleen Brady: OK.

Reva Winkler: Do you want to start with the importance to measure report criteria of impact performance gap and evidence?

Kathleen Brady: OK, yes. And so, for impact, all the reviewers ranked it as high. There was not really any specific comments I think about impact. Certainly about performance gap, there were. It was ranked high for two, two as moderate, and one is low. And the comments around that was actually, first of all that (Gardner) is not a systematic review. It is just a review. That 97 percent met

the standard in a small sample. And the question is, is there really a performance gap?

So, that's something that we really need to worry about. That there was no disparities data available, and so in the (inaudible) data that was presented are on viral suppression and not on antiretroviral therapy prescription.

In terms of evidence, for quantity, all five reviewers ranked it high and the comments were that, you know, the quantity of studies was over 45 with 20 randomized controlled trials. More than five with a large patient population including randomized controlled trials.

For quality, it was given high for four and moderate by one. The quality of evidence is moderate given and I think this might be my comment that randomized controlled trials available are mainly for persons with CD4 count of less than 350, and only cohort data is available per person with CD4 counts of 350 to 500.

Consistency was evaluated as high by all reviewers and the data – the only comment with the data is consistent in direction and magnitude.

Reva Winkler: Thanks. Any other …

Kathleen Brady: Based on the decision logic, I think everybody voted that was the scientific acceptability was yes for all five.

Reva Winkler: I think you mean important but that's good.

Kathleen Brady: Oh, yes, (important). I'm sorry. Yes, that's the one I'm talking about. I'm sorry, that's not true. (Inaudible) for importance, it was yes for four and no for one. I apologize.

Reva Winkler: OK. All right. It sounds like there might be someone on the workgroup that have some concern about how the measure meets the criteria. Any comments from other workgroup members?

Thomas Giordano: Yes, I'll start. It's Tom Giordano. I rated the opportunity for improvement as low because those are my comments about – about the reference being (Gardner) is not a systematic review. And so that's just an expert opinion. And the data presented showing a performance gap just were on bowel suppression not on – not showing a performance gap in prescription of ART.

So I think it's – it could be written differently and clearly, that would change my evaluation bit as written, I have to rate it as evidence if there's opportunity for improvement.

Kathleen Brady: And I will add, there is additional data that is not included in this. I mean, specifically, there's data that's been released by CDC that does show there are gaps in receipt of antiretroviral therapy in patients with HIV. It's in their vital signs, and it's replicating the (Gardner) cascade using surveillance data and data from the medical monitoring project.

Reva Winkler: All right. Any other thoughts or comments from the workgroup members?

Male: Is there somewhere where it more clearly outlines potent antiretroviral therapy? I thought somewhere it mentioned the CPT. Is there an actual code that encompassing of any potent antiretroviral therapy?

Reva Winkler: So, you're asking about the definition of the antiretroviral potent?

Male: Yes.

Reva Winkler: OK.

Kathleen Brady: (Inaudible) numerator statement as – but it's just described as any antiretroviral therapy that has demonstrated optimal efficacy and results in durable suppression of HIV as shown by prior clinical trials.

So, it's a fairly vague definition.

Male: So, how is that captured? I mean, is there a list of drugs and if a patient is on certain drug or do – do individual patients have that CPT code that says they're on one of those drugs and you're looking for a code?

Jeffrey Beal: Hi. This is Jeff. I can imagine that there's any code out there for that. I would assume it's just by the DHHS Guidelines, the listing of preferred regimens which (inaudible) patients.

Reva Winkler: All right. Does NCQA want to comment?

(Jenna): Yes, thanks – thanks for the opportunity. We – the way this is captured is that it's a Category II Code that the provider uses to test that the patient is on potent ART. We do refer providers to the HRSA guidelines about – that describes what treatment that patients with HIV should be on, and the reason why we took that approach is that the treatment guidelines changed quite recently and it would be very difficult to maintain the list of acceptable drugs and acceptable combinations, and be able to have a Category II Code that ties to those acceptable and recommended combinations.

So, we do use a Category II Code that's just as the potent antiretroviral therapy prescribed.

Jeffrey Beal: OK, so then – this is Jeff. My question to you back would be then do you need to make this only apply to naïve patients because that's where by far, the majority of clinical trials regarding antiretroviral therapy as in a naïve patient. We have a limited number of studies that deal with the switches after those initial events.

(Jenna): We actually discussed with our expert panel that we pulled together to discuss the measures and they think that the measure should apply to those naïve and treatment-experienced patients. Again, because we are referring patients to the treatment guidelines, they can look to the treatment guidelines to explain what treatment should be used for the experienced patients as well as the naïve patients.

Jeffrey Beal: OK.

(Aaron Milstone): So, you're not actually measuring whether patients are getting potent antiretroviral therapy. It sounds like you're measuring whether or not providers are documenting that a patient is on a potent antiretroviral therapy.

(Jenna): I think that is the more accurate way to think about it, yes.

(Aaron Milstone): And just again, in other samples – this might – you know, air out some of my naivety about how some of data has been captured. But outside of the test set, do we know how well most providers are at documenting that Category II Code versus actually putting their patients on antiretroviral therapy?

(Jenna): I'm using these particular codes. The only data we would have would be from the PQRS program. We don't have data available right now and I don't have that on my fingertips data about how providers are doing about documenting this outside of the PQRS program.

(Aaron Milstone): I'm sorry. What's that program?

(Jenna): It's the Physician Quality Reporting System program. It's run by CMS and it's an incentive program for providers to report quality measures.

(Aaron Milstone): And do you have a sense of how many health HIV care providers participated in that program?

(Jenna): We did. In the (inaudible) commission forms, we have actually provided the data about the number of providers that have reported on these measures in the years when they report it. Some of these HIV measures are included in the PQRS program, not all of them are.

So, I'm just referring to our testing section right now. It's under Section 2B 5.1 under the data sample, under Identification of Meaningful Differences.

(Aaron Milstone): Like, I'm seeing – doesn't it say on page 14 in 2009, there were 60 eligible providers reported this measure.

(Jenna): Yes. The HIV measures I will say, they were recently introduced to PQRS. They do have a relatively small number of providers reporting right now but

we would certainly hope that the measures will become more widely used in the future. It is a voluntary program.

Thomas Giordano: This is Tom Giordano. So, to follow on that – that set of question. The provider at the time of the visit has got to code this or is there a coding person behind the scenes that does this, or is this something that – I'm not familiar with the process for this.

(Jenna): I think that – sorry.

Female: No, go ahead.

(Jenna): Yes, I was just going to say I think the actual workflow is up to the individual provider. It's more likely to me that they would actually have an abstracter at the point when they are submitting codes for the reporting program, that they would have an abstracter go through and pull the data they would need in order to report the codes.

I'm not – I don't know the – I know the program well enough to be able to speak to that kind of detail.

Thomas Giordano: So, at a more general level, to me this says how feasible is this or usable is this measure and that – I know that's a criterion but I guess what's the – what's the role of this committee in trying to figure out yet how widely something like this could be done because it is time consuming to figure out if a combination that's being used in an experienced patient would qualify.

(Aaron Milstone): Exactly. And to add in, it's interesting the discussion is very helpful to me because actually, you could have a patient who is highly resistant to a highly potent regimen that maybe in this group of OK drugs to prescribe. Yet, that's not the ultimate goal of what we want. We want the patients to be undetectable for a large number of reasons.

So, separating this out from the viral load measure, now strongly makes me question the usefulness of this measure.

Female: Yes, I mean I think though it's a process, you know, along the road and not – obviously, not everyone who has prescribed potent antiretroviral therapy become undetectable. And you want to may – I mean, I think the difference is there are some people who are not being prescribed antiretroviral therapy even though they should be. And that's really what we're trying to look at in this measure.

(Aaron Milstone): I guess my – that I agree and my question that comes back to are you measuring whether patients are getting prescribed highly potent antiretroviral therapy because you're not capturing the drug. You're capturing whether someone code that a patient is on a drug.

Female: I know, I do understand. I mean, that's an important thing and I think maybe that's something that deserves maybe further evaluation to, you know, find out if the way if that's the coded, is it our people truly getting potent antiretroviral therapy. Because if you have data to show that, yes, we reviewed all these records, and you know, here are the regimens that people were on or you know, these are the ones that met the criteria as potent and these were the ones that didn't.

I think you would probably feel better about using this measure, am I right?

Jeffrey Beal: Not me. This is Jeff.

Female: No?

Jeffrey Beal: You know, I have to be a devil's advocate here and say, number one, it's going to be very hard to have that somebody abstracting and programming this measure from the standpoint of all of the options in ways that we can put things together. The key is if the patient undetectable on viral load or not. And why waste the effort on this, number one if they're not undetectable on viral load, it doesn't necessarily mean that they're not on an antiviral regimen.

But more than likely, if they are undetectable on viral load, they are on an antiretroviral regimen. So, why put the effort into this if the endpoint that we're seriously needing to be concerned about is undetectable viral load?

Male: But, Jeff, let me – let me …

Female: But there are differences in who was prescribed the antiretroviral therapy or who's not prescribed the antiretroviral therapy, and who doesn't become undetectable. There are two different types of intervention that, you know – you know, someone may not be undetectable because their doctor didn't prescribe it. But to some, it may be undetectable because they didn't take their medicine.

I mean, I think you're – it's looking at two different things.

Thomas Giordano: It's two different steps in the process, and I think there is value to measuring the steps independently as well as a global performance. But I – I mean, I respectfully disagree with you, Jeff. I think that there are some values to measuring the steps. I think alone, it's not enough. But I think it's an important step.

Jeffrey Beal: That helps. I – you know, I get that.

Thomas Giordano: Can I – this is Tom again. Can I raise another issue? I think you could have the exact same debate about the difficulty in calculating the numerators you could have (inaudible) calculating the denominator because you have to have an AIDS-defining condition or a CD4 less than 500. And with pregnancy, I mean, there are good codes for pregnancy. There are not however, ICD-9 codes. There are ICD-10 codes for AIDS versus HIV.

So I don't know how you operationalize that and again, this gets to the usability of the measure. But that's, you know, it's difficult to figure out who's got AIDS and who's got HIV just based on their ICD-9 codes. And you've got to go back to (forever) because once you've got AIDS, you've always got AIDS.

Reva Winkler: Super conversation. Anything else on – you know, we kind of gone through some usability and feasibility as well as reliability and validity but any other thoughts before we move on to another measure? Any other questions to the developer? OK.

Then, we go on to the next measure which is really addressing sort of a similar concept and that's Measure 2083 which is a new measure from HRSA which is prescription of HIV antiretroviral therapy percentage of patients regardless of age with a diagnosis of HIV prescribed antiretroviral therapy for the treatment of the HIV infection during the measurement year.

So, Sue, I think this one is yours.

Sue Elam: OK. So, I have a question and that, you want me just to go through as everyone else has done with just what the consensus was on the evaluation? It's not so much …

Reva Winkler: No, add your own thoughts.

Sue Elam: Oh, no. I just didn't know if I needed to go through that it is a new submission and basic …

Reva Winkler: Yes.

Sue Elam: OK. So, if we look at the consensus from the workgroup, we have consensus on the fact that this a high-impact – it's a high-resource use patient social consequences of poor quality and severity of illness. So, there is a consensus there. When we look for the performance gap, basically the group said that three agreed that there was high opportunity for improvement. We had one moderate and one insufficient.

Some of the comments that came with that – insufficient data to require treatment of all patients with HIV. This does not provide exclusions for people that refuse treatment or are not given treatment for various reasons. And then the – the data was somewhat limited for persons with CD4 counts over 500 – I'm sorry, that's in to the body evidence. Sorry,

Reva Winkler: That's fine.

Sue Elam: So, the comments on the importance were that it didn't show deficiencies and antiretroviral therapy prescription and the other comment was increased

treatment equals decrease viral load, equal decreased transmission, morbidity, and mortality.

Reva Winkler: Comments from other members of the workgroup?

Aaron Milstone: This is Aaron. I'll start. I was the one who had the insufficient for the quantity and I – and I again, I think this is important. I think we all recognize the importance of this clinically but just as a pediatrician, I just wanted to point out that you know, the current guidelines that are presented for pediatric guidance in children less than 5 years of age, it says; for those that are asymptomatic with a CD4 percentage rate of 25 percent or an HIV RNA less than a hundred thousand copies, that you should consider treatment.

So, if you have a group in which the recommendation nationally is to consider treatment. This does not in any way taken as a new account. This is all patients. So, I'm trying to – I was trying to settle with myself how you could say that we have a measure that applies to everyone when the current national standard is that some people shouldn't be treated or don't need – or there's no recommendation for treatment in all patients.

Kathleen Brady: It's Kathleen. Do you think the measure should be modified then to say that for a person less than 13 years of age?

Reva Winkler: Marlene from HRSA. Did you want to comment?

Marlene Matosky: Sure, thank you for that opportunity. So, my name is Marlene Matosky. I work at HRSA's HIV Bureau and I just wanted to start by saying, when we were first introduced although that we're listed as the developer in the measure forms, I just wanted to note that this was a project that had a significant and excellent partnership with CDC and John Brooks and Abigail Viall are also on the phone. So, at any point, they can also provide comments on this also because they're part of our steward group.

But going back to what Dr. Aaron had talked about is that you know, first and foremost, we see this as not just a clinic of facility measure. We're introducing this as you know, being able to be utilizing the clinic but also looking at it as a population measure.

So, this could be used for – within the region and within the state. And so, we not included the aspects and the consideration of the guidelines that I'd go towards when to treat but we also took into consideration the aspects that are related to prevention and all these theme that's in order to recently, and you know, somewhat, you know, also promulgated by the international conference around treatment as prevention.

So, that's one thing I would want to put in there as a comment. Another thing that comes to mind when we look at these is that we at HRSA and you know, our colleagues at CDC, we've never thought performance measurement as being a zero or a 100 percent. We never have set those as the patient's benchmarks or goals.

So, therefore we know that there are going to be some folks who are not going to make it into the numerator, and you know, following from that, we see performance measurements as just one side of the quality improvement coin. The other side is that quality improvement piece or I'm sorry, the quality measurement coin, the other side is the quality improvement.

So, you know, we're not just utilizing these for performance measurement. Our ultimately goal is for that quality improvement piece also. So, we would feel as though it would be expectation that if you're utilizing this measure, you would also look to see who's not making it into the numerator, determine if there are any trends, and decide where you go from there.

And you know, you being a pediatrician, probably could speak to this. You know, we know that we have a very stringent review of our pediatric population and that there is always consistent monitoring, and even more in-depth monitoring on that population to see where they are in their HIV disease and seeing if they're eligible to be put on to antiretroviral therapy.

So, I hope you know, thinking about one – you know, considering, you know, treatment as prevention comments, thinking about it as much as the performance measurement that it's also the quality improvement. And then thinking about, you know, the increased monitoring of the pediatric

population and knowing that we're never a zero or a 100 percent with anything.

I did invite colleagues from CDC to add anything as they'd wish.

John Brooks: Thanks, Marlene. This is John Brooks. And I just – I would agree with everything you said. I did wanted to also add a brief comment about the issue around limiting the measure to persons whose CD4 cell count is greater than 500. And I say this in part and a participant in the – in full disclosure. I'm on the committee that helps write these guidelines. But I can't disclose the committee's deliberations but what I can say is the intent of the recommendations for – the (inaudible) recommendation for treating over 500 is not so much to say that you can't do it.

The "consider" really is intended to interpret as "may treat" that everyone is permitted to treat over that CD4 cell count. And in fact, that we know that some large jurisdiction including San Francisco, New York City, and I believe Washington D.C. now are making a policy that all patients diagnosed with HIV infection regardless of CD4 cell count are being treated.

So, if we apply this limitation, I just would question about its – the utility because we may be excluding large fractions of people who'll be increasingly treated. It's true that most patients who are identified with HIV infection in the United States, they are identified at a CD4 cell count under 500. But if we at our agency and as others, do our job right, we're going to – we hope that that's going to – that more and more it will be the case that we identify people very early in the course of their disease.

So, I guess that's really what I wanted to say there. I should also note that some of the large systems – I know that the (A-system) now permits treating any patient with HIV infection as policy and I'm not sure what Kaiser's policy is. But I think, you know, we're seeing that a lot of jurisdictions who – for whom the public health perspective is very important are really pushing this.

Reva Winkler: Any other thoughts from the workgroup?

Thomas Giordano: This is Tom. Yes, so, thank you both for those responses. And I get the spirit of what you're saying and I agree with you completely. But is that the goal of the quality indicator? You know, is it – is it – what we were tasked with was to say does this – is this measure based in the science and is it something that can be used? And I think as clinicians, as people who care about the public health and care about our patient with HIV, we will do our best to treat the patient and that may mean treating everyone regardless of CD4 count.

But as people who work in clinics where there are administrators, we don't want to say you're going to be held to a standard that isn't backed up by the data. You know, the last thing we want is to have a new measure to come in and say, well you're not – here's your data and you're not meeting the standard and they don't understand this because you've got a large pediatric population. And they don't understand because you've got – you're in an early intervention clinic and you're seeing a lot of people with a CD4 over 500.

So, the spirit, yes, I'm on board, completely. But you know, I don't know how we're supposed to handle that given the rules that were set up for us by the – by the people who charged us with this task.

John Brooks: Tom, this John Brooks. Thanks. It's a very good comment. I hear you. I understand what you're saying. I've often struggled with the same thought, and I guess it depends on how – in terms of quality, if the quality measure whether the science and also the official U.S. recommendation would say that it's poor quality, not to treat over 500. And that's what I – that's the kind of way I've been thinking about it.

And I would say that, you know, you're not going to be penalized for treating over 500, and that is – it is backed up by recommendations in many place that treat are taking that part and following that as a recommendation or a reason for them to not treat at diagnosis.

So, I'm not. I feel that if you're going to – if you don't include that you're going to be missing a lot of possibilities for measuring other quality

performance that people are engaged in. And also what leaves you more flexible into the future in case, you know, practice begins to change for which the evidence is growing that is beginning to move – beginning to move in that direction.

Male: So but – then aren't you suggesting that the quality measure will drive practice? I mean, I come from the background in infection control where lots of things that are reported for quality improvement has – are now not done for quality improvement purposes. They're done for reporting and incentive based compensations, et cetera.

So I think if you think that treating people with CD4 counts over 500 is important, then that should be evidence based, but practitioners shouldn't do that so that they are faced with criticisms for having a lower than expected reflection in their quality measure. And Tom, is that – I think maybe we're saying something similar.

Thomas Giordano: Yes.

Male: Yes, and I'm curious about other people's – how, where we come into this conflict. Because my concern is putting it for the quality measure that drives a practice that isn't – that we all agree maybe worthwhile but isn't supported by evidence.

Reva Winkler: OK. Any more thoughts from the committee? Perhaps we can talk about reliability and validity. Sue?

Sue Elam: OK. So, on reliability, the group rated one as being high, four at moderate, and the validity was four at moderate and one at low. And the comments were that it was a good sampling of patients. The testing details were explained. What about exceptions that are not accounted for and then, the concern about the validity – not concern but noted that it was face validity and so that, I guess would justify the moderate rating, the highest rating being moderate under those circumstances. And then there was a comment about questions of threats to the validity.

Reva Winkler: Comments from other workgroup members? So, I'm hearing none and that – I'm concluding that the workgroup really does not have any concerns about the reliability and the validity of the measure as written. There is a relationship between evidence and concept validity. So, but if there are no comments, well…

Male: Well, I think what we said before would apply that – that and I think the comment there addressed it that if you're including people with a week per indication, with a CD4 count greater than 500, then is that – does that mean there is a less validity in measurement.

Reva Winkler: All right. Any thoughts on usability and feasibility?

Sue Elam: The group had consensus with three being at high usability, one at moderate, and one at insufficient. Comments included that it was meaningful, understandable, and useful to the intended audience, public reporting, and quality improvement. And then, process for reporting is not outlined as far as the inclusion is pending with the CMS. Any discussion on those?

Female: I did notice that comment under feasibility that someone said that the list of interviews has some potentials for difficulty in data collection. Could you sort of explain that a little bit? Whoever? That doesn't ring a bell to anybody?

Male: I don't remember if I wrote that comment or not. But actually, I like this approach of at least outlining the medications that should not be used together better than the approach of saying someone is going to try to review the regimens for the other measure and say that it's consistent with or inconsistent with the guideline.

But it is difficult to imagine someone doing this on all patients. I mean, I suppose you could write a computer program that would do this, if you've got a sophisticated enough database.

Female: What page is that described on?

Male: Twenty-four.

Female: OK. Thank you.

Reva Winkler: Any other comments on this measure before we move on to the next one? Any other questions for the developers? All right. Clearly, we do have two measures that address similar process of care around prescribing. And so, after review of the individual measures, there will – we will want you to kind of take a look and see the two different approaches for measuring this process of care, and see how you feel about them in terms of meeting the criteria and usefulness of each of them. And we'll do that at the in-person meeting.

So, that brings us to our next measure which is 407 which is HIV-RNA control after six months of potent antiretroviral therapy, and this is for patients age 13 years and older who had at least two medical visits during the year with 90 days between, and is using antiretroviral therapy who's has a viral load less than 200 copies.

So, this measure is, Tom, I think this one's yours?

Thomas Giordano: Right. so, this is a maintenance measure that was originally endorsed in 2008 but the outline with the measure is, in terms of importance, there was consensus that it was – that it was important. Impact was rated at zero. I'm not sure what happened there.

Reva Winkler: Yes. That was four before.

Thomas Giordano: OK. I believe you on that. in terms of performance gap, there was supposed to be either high or moderate level of performance gap. It was noted that there was no disparities data available but there was significant gap in performance. One person made note that there was a single source of data on the performance gap.

All these are zeroed out.

Reva Winkler: Oh, OK.

Female: When we got the second set yesterday, the revised, I noticed that – that on this one, there were quite a few zeroes.

Reva Winkler: Yes, something obviously tripped.

Female: Yes.

Reva Winkler: So, well, I actually am sitting in front of the first version so I may be missing one. But in general …

Thomas Giordano: I've actually got the first version if you want me to…

Reva Winkler: Great. Well, you can use that as a guidance. I think you're missing one person's input but …

Thomas Giordano: Yes, that one person is me.

Reva Winkler: OK, good. You can interpolate.

Thomas Giordano: So, evidence was strong. We have two and two – I'm sorry. let me get reorganized here. Anyways, that was – the comments that in terms of the quantity of the evidence, four people felt it was high and I believe I would add a fifth vote to that. the quality was two high and two moderate, and I would rate it as high and the consistency was three high, one moderate, and I was also moderate. So, three high and two moderate.

Some of the comments there – there was a comment of why this would not include children between the age of 13 and I guess, 21. There is that in the denominator, so I'm not sure what that comment means to be honest. If someone could clarify that maybe?

Female: (Inaudible) was that supposed to be less 13?

Male: Yes, that was mine. That was less than 13. Sorry.

Thomas Giordano: Do we want to discuss that or …

Reva Winkler: What are your thoughts, Tom?

Thomas Giordano: Well, I mean, I think for the reasons you've already pointed out, there is a different standard in – in children that is way outside my area of expertise. But, so does – would this apply with the supply for children in that age range?

Kathleen Brady: I mean, I guess my questions is, why wouldn't it? this is Kathleen again. And I would think that if you start a child on any antiretroviral therapy, you know, who's less than 13, that you would expect them to be undetectable by six months. Is that correct?

If there's a different standard, then that's why…

Thomas Giordano: Oh, no. I mean, that's why I asked why they weren't – I mean, this isn't are you starting or are you not? But if you do.

Female: If you do…

Thomas Giordano: Why wouldn't they have the same standards?

Kathleen Brady: Yes. So, if it's the same standard, then there isn't any reason why it should be – it shouldn't be everyone.

Reva Winkler: Comments from the developer, from NCQA?

(Jenna): I think that's a really interesting point, and I would have to discuss that with our expert panel about whether they would be comfortable with lowering the age limit. My suspicion is that one reason why it might have been developed this way is that it's not necessarily a paired measure but it's a link to our other potent ART measure and the denominator for that population is greater than 13.

So, I think we were just trying to align denominators. But I do really appreciate the comments that if – once the patient has started that the treatment goals might be the same. So, like I've said, happy to take it back to our experts.

Thomas Giordano: Any other comments on the importance before we move on to scientific acceptability?

Reva Winkler: OK.

Thomas Giordano: So, moving on to scientific acceptability then. for reliability, the ratings were one high, three moderate, and then I actually entered it as low for reliability. And my rationale for that was because the measure as written to me was very ambiguous as to which viral load you're supposed to use. Is it after six months of potent ART, you could've more than one viral load in the 12-month measurement period? So, do you use the last viral load? Do you use any viral load? Do you use the lowest viral load?

So, I thought there was some ambiguity there that needed to be clarified in the – in the description of the measure.

Kathleen Brady: I would agree with that. this is Kathleen.

Reva Winkler: Anything from you, (Jenna)?

(Jenna): I think that's a good point and again, happy to work with our expert panel to clarify that.

Thomas Giordano: Other comments on that topic or on the reliability topic. I think similar things have already been brought up about EHR testing at the measure level. Only on the data element level, lest the validity not addressed. I'm not sure what that means exactly. And unclear how this would perform on the other side using the different EMRs.

And then again, the same question about potent and how that's going to be defined. Anyone would want to comment on those comments?

(Jeffrey Beal): I'll just reiterate what I said before yet again this is within the Category II Codes for patients on therapy. So, it would be the same discussion we had before about how has that been – and I actually didn't look at the last one to see if that uses the same method for defining patients on antiretroviral therapy but I would bring that up again on this one. How well those Category II Codes are for identifying these patients?

Thomas Giordano: So, in terms of the validity, the ratings were high and moderate, two and two. And I also gave it a moderate. Any other comments on scientific acceptability before I move on to usability?

So, on usability, the ratings were three high and one moderate and I gave it a high as well. comments were (inaudible) current use CMS, so it is currently in use and shows need for improvement. Any comments on usability?

Then feasibility, high it got to votes, and moderate it got two votes. I gave it a high overall, I believe. Oh no, I'm sorry. I did not. I gave it a moderate overall. Again, feasibility to me was limited by simple issues. One was the issue of bleeps. I mean, you are using a definition of 200 as bowel suppression but that is not an absolute definition for a bleep.

You have to set a line somewhere, I understand but you could have bleeps that go higher than 200. And then, I think it's going to be difficult to know if someone was on ART for at least six months. I mean, I guess there – I guess I learned today that there are codes for someone being on potent ART but just in terms of real patient care, saying that someone was – understanding what the patients are on therapy and when they're off therapy because of the side effects or for whatever or other reasons.

It's not easy and just managing this measurement standpoint, I see it may be difficult to operationalize on ART for at least six months.

And then of course, there's going to be a lot of missing data. I don't know. None of those measures really address how you handle missing data but…

Any other comments on the – on the feasibility.

Female: I would just agree with what you said about the operational part of it with regard to ARTs and things that can happen with viral loads for a variety of reasons.

Female: And especially, I think the six month issue is definitely something that maybe difficult to operationalize.

Reva Winkler: This is Reva. I just wanted to ask if NCQA wanted to comment on how they handle missing data.

(Jenna): I can because this is Physician Reporting Program. we don't necessarily spell out in the measure how they should handle missing data. I'm just trying to think through this. I apologize. So, when you say missing data, does that mean missing and not having the lab value?

Male: Exactly.

(Jenna): Right. So, if it means that then, if they don't have a lab value less than 200, they would actually be a numerator miss if they don't have information about their RNA labs. Because we don't allow – there is not an exclusion for missing lab values. They obviously can't be under 200 if they don't have the lab values.

Male: But I think that's the perfect way to handle it. Great. Thanks for that clarification.

Reva Winkler: OK. Any other questions on Measure 407 from anybody in the workgroup or anywhere? OK, well, we also have a similar topic in the new Measure 22 from HRSA. This is a new submission that HIV viral load suppression that's presented to patients regardless of age with the diagnosis of HIV with the viral load less than 200 at last HIV viral load test during the measurement year.

So, Jeffrey Beal. I think this one's your measure?

Jeffrey Beal: Yes, it is. Thanks. And just another patient exclusions in this measure either. So, the group under impact was unanimous in rating it as high. It was supported by clinical trial evidence so then antiretroviral therapy reducing the morbidity and mortality of HIV as well as improving the quality of life.

And also the emerging evidence that's coming about antiretroviral therapy decreasing the chronic inflation that is resulting in some of the complications and also in reducing transmission of HIV.

As far as the quantity, we were three with – and I guess we might have an – three with high and one is moderate and quality three as high and two as moderate and consistency three as high, and two as moderate.

With quantity there were six randomized controlled trials, one being a meta-analysis, eight observational studies, and the reliability testing was done through the multi-site HIV research network which is inclusive of both community-based centers and academic centers providing HIV care, representing four major geographic regions.

And the analysis of this data is by face validity with a technical workgroup and then through a webinar done with national representatives of Ryan White providers. And we basically rated this reliability and validity as moderate.

Reva Winkler: Any thoughts from other workgroup members? Looking at some of the issue you raised with the other measure this one does define which test is captured. This does capture children as well. Any comments on that?

Aaron Milstone: This is Aaron. I have the same comments here that I did before. I recognize that there's a movement toward treating all children with HIV but currently, there are providers that don't treat children that are asymptomatic high viral load and high CD4 counts. So, this doesn't seem to account for that.

So, that was why it was insufficient. I just thought it was insufficient in that population which is part of the overall population. And I do – we can have this in discussion. I do take to heart the previous comments. I thought they were very nicely laid out by CDC and other.

Reva Winkler: Go ahead.

Thomas Giordano: It's Tom. The viral suppression measure, I guess, it's sort of a snapshot of the whole process. so, they have to be on ART and here and I like that because it's – you know, that is the end result of our efforts is to keep people suppressed for a lot of different reasons.

But when you don't look at each step in the process, you ignore or you can miss important reasons why someone might not be suppresses. Maybe they

have refused ART, or maybe they are not adherent for whatever reason, maybe they're having – they have other issues that are actually more important from a clinical standpoint.

So, does that matter from a measurement standpoint? Does it – if the measurement is so – is really great for public health but may not make a whole lot of sense at the individual level, do you need to put some restrictions on the population in the – for the denominator? I don't know the answer. I'm just raising it as a question.

Kathleen Brady: It's Kathleen. I mean, I think in general with any indicator, the expectation I think like Marlene said, it's never going to be a 100 percent. And so, I mean I think that's what you have to keep in mind, it's what the benchmark that you know, we're looking at.

Marlene Matosky: Reva, this is Marlene. May I make a comment?

Reva Winkler: Oh, please do.

Marlene Matosky: So, my next comment, you know, aligns with, you know, I would say probably all of our measures and I would think NCQA would also agree. The Office of the National Coordinator for Health I.T. those are the folks who are you know, co-implementing the meaningful use of electronic health record programs with CMS.

They've been you know, gathering expert testimony from a variety of different sources including the provider community and the you know, EHR test community. And we've been hearing more and more of that. We need simplified measures. And I'm not saying let's have simplified measures, you know, to have simplified measures to satisfy that group but thinking about you know, our ultimate angle here is to, you know, ultimately get endorsed.

The next step is we will be responsible to specify these measures for use in electronic health record, and then if you're using a certified EHR product or actually EHR modular product, either these – the vendors would make these measures available in your EHR or that certified module.

So, with that being said, you know, we've really drawn back from you know, adding a lot of exclusions that are not based in structured data within an electronic health record. You know, for instance, and I think it's noted in our measure submission, when we initially had you know, utilize this measure in another project, we had exclusions related to folks who are incarcerated. And you know, we soon came to realize, incarceration although it is unfortunately, an impact, has a significant impact on you know, our patients who are infected with HIV, it is not something that is you know, structured data within electronic health records.

So, it would be putting out an exclusion that ultimately, we could never you know, account for. So, I hear what you're saying. And so, we're trying to achieve that balance and then, you know, I'll just go back to a comment I made earlier about that ideas you know, balancing the measurement with the improvement piece and you know, looking at your numerator to determine trends and so on.

So, that just kind of gives you a little bit of thinking about we have come from when putting forward these items in the process. thank you.

Reva Winkler: Any additional thoughts or comments from the committee about this measure? So, everybody's got their questions answered and feel they know what they need to. All right, then we'll go to the last measure which is Measure 412, hepatitis B vaccination. This is a measures that's previously endorsed from NCQA. It is for patients aged 6 months and older with a diagnosis of HIV/AIDS who've received at least one hepatitis B vaccination or who have documented immunity.

So, this I believe, Aaron? This one's yours?

Aaron Milstone: Thanks. So, this was – I'll try to start with the impact. I mean, if people can see – three people rated this as high and one is moderate. And then in the performance gap, there was one insufficient that was mine that I'll comment on. But there were questions about the number of people surveyed, or whether there is room for improvement, no disparities which have been similar.

I guess my – the reason I had insufficient was because if I read this right, the data that's described – again, the intent of this is which I think we all agree is very important is to – is to give hepatitis B vaccination in patients with HIV. But the – and the data supports how many patients with HIV have gotten hepatitis B vaccine series.

But I did not see data on how many people have gotten one vaccine. So, I wasn't sure whether there was – and correct me if I'm wrong, where there's evidence for that. that was I listed it as insufficient in the performance gap.

Jeffrey Beal: This is Jeff. I don't think there is any evidence for that in the HIV infected. But in the normal host, there's evidence of good response to one hepatitis B vaccination in the literature that got the vaccine originally endorsed and approved. I think it was – I don't know if I remember but I think it was better than 80 percent of people seem to get an antibody response. But there's certainly no data like that in the HIV infected population. So, I think that's a good point.

Thomas Giordano: This is Tom. I had another issue is the performance gap. As I read Section 1B2, the data to show that there is a performance gap, that wasn't noted in the HIV infected population that was …

Female: (Inaudible) in that risk.

Aaron Milstone: It was in the NIHS population which is general adult population, right?

Female: Yes, I agree with that. that was actually the issues that I have with this was that the data presented doesn't actually reflect the problem.

Aaron Milstone: So, I rated it as low evidence for performance gap.

Reva Winkler: All right.

Aaron Milstone: I'm sorry. I just want to keep going and then moving on from there was the evidence itself, and I think there was consistency most people rated the quality of data as either high or moderate as well as the quality and the consistency.

I'm sorry, I'm looking at the – I'm looking at the old one. So, let me look online and make sure or correct me if I said something wrong. Yet, nothing's changed, OK?

So, I'm looking at the form from yesterday but it looks like the (yours) in relative agreement amongst the group in the evidence in all populations with patients with HIV. So, I don't think there was any question about the evidence since we moved on – I mean, one caveat that I noted again was this was based on the need for hepatitis B vaccine. It wasn't on one dose of vaccine as a measure.

I still think there are some question there which I'll come back to later. But again, the goal of it, I think the intent of this is to get people hepatitis B series not just one injection of hepatitis B. and I realized the importance for this as a marker but how effective one, and it was brought up again, how effective one dose of hepatitis B is in this population is unknown.

And then we move on to the scientific acceptability and when we look at reliability, it looks like there were, my screen hasn't moved down. Can somebody on the webinar just push it down a little? There you go.

So, it looks like there's a – OK. It looks like there were two people reporting moderate for reliability, a couple or two low, and one insufficient. And same with validity, there were concerns about the validity. And this I think based – I think a big point that was noticed here was that the discrepancy between the tested sample where the automated calculation of performance and the manual calculation performance had a pretty significant difference in percentage of up to 60 percent.

Obviously I think raise concerns by the number of the reviewers. Not sure if anyone wants to add comment to that.

Female: Yes, I was one of those people. I just think that that's at of an acceptable range.

Female: And I agree.

Male: Could the developers comment on the one vaccination bar that was set?

Female: Absolutely. I'm happy to. There are several reasons why we did – wanted this to be first shots. One reason is that if we were to think about capturing all three rather than just the first, we would actually have to change the denominator to I think allow a longer look back period since there is that timing about the amount of time between the different injections. And we also had some concerns about patient that might leave at provider's practice while they're getting those three injections. We certainly agree that the goal is this full series.

Another reason is that we do try to keep in mind harmonization and there is another – there's an AMA PCPI on the measure I believe for the hepatitis C population that has a similar construct with the one – at least one injection.

And the third comment I would make is that we discussed this with our expert panel and we settled on that the first injection does at least provide data that the provider is starting to initiate proper care. But again, because there is that time in a roll between the first and third injection and patient may leave the practice that we thought we'd like to focus on the first injection.

Male: Can you comment a little more about the – you mentioned again some category 2 codes in your numerator. And do you have any information about how often providers are using the patient has documented immunity to hepatitis B versus the others?

And a follow-up to that also, are these CPT procedure codes, are these – that the vaccine was ordered or that the vaccine was administered?

Female: The category 2, because they are actually described under numerator details so that's two-way 1.3. So we're either looking for category 2 code that describe that the hepatitis B vaccine injection was administered or previously received. There's another code for patient has documented immunity to hepatitis B and another code for hepatitis B vaccine series previously received.

Does that answer your questions or there's still an additional question you have?

Male: Well, I guess my question on that was the same as we talked that before which is how common or do people, have those codes been – I'm wondering whether or not that was part of the discrepancy in the automated versus manual where people were documenting in their chart that a patient had been previously vaccinated against hepatitis B but they weren't getting the category 2 code of 4157-F.

So I was curious if you did on how commonly those codes are being used versus the practice of getting that – of documenting somewhere else in the chart to the vaccine.

Female: Unfortunately, this one, this particular measure is not used in PQRS and I will admit the category 2 codes are mostly used for the PQRS program. Some systems or some groups might choose to use category 2 codes to do a quality measurement, but the use of category 2 code is mostly driven by PQRS. And because this measure is not included in PQRS, we don't have data about how frequently those codes are being used.

Male: So then how do you capture documented immunity? Because the measure is written as have received at least one vaccine or has documented immunity and if I'm reading this is the only way you would take out documented immunity in a numerator is if a patient has a category 2 code of documented immunity.

Female: So I would think and others please feel free to correct me if I'm wrong that documented immunity might be in a paper medical record as a note somewhere.

Kathleen Brady: I think – it's Kathleen. I think most commonly it's going to be in laboratory results with evidence of someone having a hefty surface antibody actually. And it's actually recommended that that be checked in persons who've received hepatitis B vaccination to make sure that they've been adequately vaccinated and people with HIV. So it seems to me a better marker looking for hefty vaccination might be to actually look for evidence of immunity through laboratory results.

Thomas Giordano: This is Tom. I would follow that up with I agree with Kathleen completely. This gets to feasibility as well as reliability. The way the measures are written the numerator I'm not sure it reflects clinical reality. You don't do this every 12 months.

Female: Right.

Thomas Giordano: You do this once. And if a person has been an HIV carrier for a decade, they got their (HBV) vaccine series 10 years ago and no one is going to keep indicating every 12 months that the person is immune. They're not going to do a lab test every 12 months for it, they're not going to necessarily chart that anymore. So I have real questions about how you can even – how you can measure this if it's supposed to reflect clinical reality. But not if you're expecting people to document this every 12 months then the measure has problems because that's not clinical reality either.

Female: I mean, I think it kind of depends on how it's documented in medical records, you know, and the one that I use there, you know, if someone has HIV automatically hep B vaccination comes up in a maintenance of health field and you can get rid of it by, you know, putting information in regarding the fact that they have a hefty surface antibody that's positive.

What I mean, I know that's all coded differently, et cetera, in different medical record systems. So I mean, I think that goes to, you know, how you can operationalize this.

Thomas Giordano: Right. I just see a lot of ambiguity in the way it's written about whether you're expecting us to be addressed every 12 months or within the 12-month measurement period or whether you're saying it has to be addressed at least once since the person entered HIV care.

Male: Now those are excellent comments. The only way I can see that you can fix this measure to keep it would be to say new patients who have been in care for the past 18 months to may be 24 months you'll find this information readily available.

Having done a ton of record reviews in electronic and also in paper through the State of Florida with Florida (Inaudible) (AAPC), this is the hardest thing for us to find because we have to go back to old, old, old volumes and nobody ever puts it in the new medical record, just practically speaking.

Female: Certainly I appreciate those points and that's something that we can think about and discuss here.

Male: One more follow-up on the numerator. Again, the CPT procedure codes, are those looking for a vaccine that was ordered or a vaccine that was administered?

Female: It is administered actually, I believe. I was just – administered or previously received.

Male: OK.

Female: All right. So …

Male: I think that summarizes the (inaudible) about reliability and validity. I guess when can move onto the usability. It looks like from usability, there were three moderates and two lows and the reported concerns were – that there is no expected date is noted for disclosure. I'm not sure who mentioned that.

Any maybe other comments about this? I think we've already talked about so many concerns have come up already with usability, given the questions about the way the measures are going to be collected at the numerator level.

Female: All right. Any other comment on this measure from any of the other workgroup members or any other members of the committee that might be on the call?

Male: Can I just bring up one other question?

Female: Please do.

Male: You touched on – so I know in reading the packet, there was some question – even with – even amongst the measures work group about setting the bar in

one vaccine versus three and I know that you said that there were – you were concerned about the timing of that. I turned that – some of the kind of adult providers to say – I mean, in your practice, how often are you actually giving the vaccine to new patients.

I mean how, if given the small number, then maybe it is worth taking a longer look in trying to capture three vaccines versus one. It sounds like most people have it done – most people have been in the system long enough that they already have this done and then you're really looking at new people, which is going to be a small numerator and a small denominator. Because wasn't there a proposal, they consider changing this to people that are more newly diagnosed with HIV.

Male: I think that makes sense. I mean that's – in our clinic, that's what we do. We only look at this with people who recently entered our clinic, whether they're new diagnosis or not, it should be assessed for everyone as they enter the clinic. So we don't have to go back to the paper records (inaudible) and try to figure that out. But that's up to the developer. It is certainly something that we should and do offer to all our patients, the vaccine, who we don't have evidence of prior vaccination or immunity already or active disease?

Reva Winkler: Yes. This is Reva. You're absolutely right and that your feedback is going back to the developer and we're not really here to change anything per se but to offer recommendations and I, you know, we're asking the committee to evaluate and make recommendations on the measures as submitted even though you may have wonderful ideas, we're making them better.

So are there any more – we've got a few minutes, are there any more questions or comments on any of the measures here that we discussed today, from any of the workgroup members or any other committee members?

OK. What we're going to do at the in-person meeting is something very similar to what we've done today. Each of you has been the lead on one or two measures. We'll be going through them, you know, the agenda will lay out these measures in order and we will be going through them with the entire committee and we will be providing the updated summaries of this

conversation into a format very similar to the summary you're working off of today only you will get it for all the measures in the project.

As a lead, we would ask you to pretty much present like you've done with the summary but for each section, the entire committee is going to vote and we have an electronic mechanism for voting. So we'll have you vote on impact opportunity evident and reliability validity, usability, feasibility and then do you feel the measure meets all of the evidence about endorsement criteria as laid out by NQF.

So that's what we're going to be doing in the in-person meeting. Today's call was an opportunity for you to share your initial thoughts, ask any questions or clarifications from the developer. They can hear what your initial thoughts are. If there's additional information, they may want to provide to you. That would help clarify some things. That would be perfectly fine. But otherwise, essentially, this was a first patch review to give you a chance to kind of shake things through before we go to the in-person meeting where your votes and decisions are pretty much the final assessment.

Does anybody have any questions about what the next steps are and what to expect for the in-person meeting?

Thomas Giordano: This is Tom. Just a quick thing. Do the developers have a chance to revise anything based on the feedback from today's call?

Reva Winkler: Yes. We certainly can let them do that. If they should desire, we'll have them open it up and we can have revised submission forms posted to SharePoint prior to the meeting, you know, if they can get that done by the end of the next week. But that opportunity is available for them, yes.

Any other questions or comments from anyone? All right, operator, is there anyone in the listening audience who might want to ask a question for public comment?

Operator: If you would like to ask a question, please press star one.

Again, that was star one for question.

Reva Winkler: There are no questions out there? OK.

All the calls in our in-person meeting our public meetings and so we do offer folks who may be listening an opportunity to ask. So we've got a few minutes extra time, which I just want to be sure that there aren't any outstanding questions or concerns from anyone.

Alexis, Adeela and I are available for any questions that you may have between now and the steering committee meeting, please don't hesitate to contact us and let us help you with your questions.

Otherwise, if there's nothing else from anybody, I want to thank you all very, very much for the time you've already put in, in looking at these measures. The discussion today was very robust and fascinating actually from my perspective and I've done this for a long time, so I really thank you for the thoughtfulness and the detail that you're really thinking about these measures. It's really a fun to listen to you (inaudible) but unless there are any other questions from anybody, I think we're finished and I wish you all a very good day.

Male: Thank you very much.

Female: Thank you.

Reva Winkler: Thank you all.

Female: Thank you.

Female: All right. Bye-bye.

END