NATIONAL QUALITY FORUM

+ + + + +

INFECTIOUS DISEASE ENDORSEMENT MAINTENANCE STEERING COMMITTEE

+ + + + +

WEDNESDAY
AUGUST 29, 2012

+ + + + +

The Steering Committee met at the National Quality Forum, 9th Floor Conference Room, 1030 15th Street, NW., Washington, D.C., at 8:30 a.m., Steven Brotman and Edward Septimus, Co-Chairs, presiding.

PRESENT:

STEVEN BROTMAN, M.D., J.D., Advanced Medical Technology, Co-Chair
EDWARD SEPTIMUS, M.D., FACP, FIDSA, FSHEA, HCA Healthcare System, Co-Chair
JEFFREY BEAL, M.D., AAHIVS (via telephone)
MARY BLANK, MPH, CIC, CPHQ, Highmark, Inc.
KATHLEEN BRADY, M.D., Philadelphia Department of Public Health
DOUG CAMPOS-OUTCALT, M.D., MPA, University of Arizona, Phoenix
CURTIS COLLINS, PharmD, MS, BCPS, University of Michigan Health System
SUE ELAM, BSN, PHN, MHS, FNP, Kaiser Permanente Medical Group
MOHAMAD FAKIH, M.D., MPH, St. John Hospital and Medical Center
MICHAEL C. FARBER, M.D., Department of Vermont Health Access
THOMAS M. FILE, JR., M.D., Msc, MACP, FIDSA
THOMAS GIORDANO, M.D., MPH, Harris County Hospital District
PETER HAVENS, M.D., MS

AARON MILSTONE, M.D., MHS, Johns Hopkins Hospital
REKHA MURTHY, M.D., FRCP, FACP, Cedars Sinai Medical Center
TIFFANY OSBORN, M.D., MPH, FACEP, Washington University/Barnes-Jewish Hospital
KALPANA RAMIAH, DrPH, MPH, Msc, CHES, CPH, CTTS, American Institutes for Research
DAVID SPACH, M.D., Harborview Medical Center
ADAM THOMPSON, Consulting

NQF STAFF:

HEIDI BOSSLEY
HELEN BURSTIN
ADEELA KAHN
NICOLE McELVEEN
ALEXIS MORGAN
KAREN PACE (via telephone)
REVA WINKLER

ALSO PRESENT:

JOHN BROOKS, Centers for Disease Control and Prevention
LAURA CHEEVER, Health Resources and Services Administration
KERI CHRISTENSEN, AMA-PCPI
DIANE JACOBSEN, Institute for Healthcare Improvement
MARLENE MATOSKY, Health Resources and Services\Administration
MARJORIE RALLINS, AMA-PCPI
BOB REHM, National Committee for Quality Assurance
EMANUEL RIVERS, Henry Ford Health System (via telephone)
ABIGAIL VIALL, Centers for Disease Control and Prevention
JENNA WILLIAMS-BADER, National Committee for Quality Assurance

C-O-N-T-E-N-T-S

P-R-O-C-E-E-D-I-N-G-S

(8:01 a.m.)

CHAIR SEPTIMUS: Good morning, everybody. We are going to sort of start where we left off yesterday but I just want a slight addition to the agenda. As you know we didn't have all the documents when we talked about reliability for sepsis.

There are documents that were circulated at the end of the meeting and other documents that will be given to the committee a little bit later this morning. We thought the best time to re-discuss that would be at a working lunch. So when we have our lunch break we'll spend part of that time looking at the reliability based on the new documents that we had. Otherwise I don't have any comments about yesterday. I don't know if Reva or Helen have anything they would like to add.

MS. WINKLER: Thank you all for your perseverance and stamina. So, the question I would ask of you if you have any

questions. Clearly we do have to squeeze in these three measures from yesterday into our evaluation time frame today. I think everybody's pretty familiar with the process and what the expectations are so I think we're all mindful of time, realizing that by about 2:30 or so I'm expecting to see people having to leave. So we would want to try and get the bulk of the work done before everybody starts having to leave. So, if you've got any questions about how things went yesterday feel free but otherwise I think we could all work together to efficiently get the work done today.

CHAIR SEPTIMUS: Okay. Also, we're going to start with, as I mentioned, 0412 until Diane gets on the phone to talk about the central bundle compliance. So Aaron, you're up for that. But first let's see if NCQA has any comments they want to make as a developer and then Aaron, your comments.

MS. WILLIAMS-BADER: Great, thank

you. Hi, my name is Jenna Williams-Bader. I'm assistant director for performance measurement at NCQA.

Today we are presenting a suite of eight HIV measures to you. We're going to be talking about I think just a couple this morning.

The measures were originally developed in 2008. It was a collaboration between NCQA, the AMA-PCPI, HRSA and the Infectious Diseases Society of America HIV Medicine Association.

We did pull together an expert panel for the creation of those measures. It was a multidisciplinary panel and you'll see the list of those panel members in your book.

The measures were originally created to be used in the PQRS program which you heard a lot about yesterday. So they were originally specified with category 2 codes. The measures were tested, received time-limited endorsement from NQF and

underwent testing in the EHR similar to the process you heard AMA-PCPI describe yesterday for their hepatitis C measures.

And the reason why we tested in the EHR similar to the reason the AMA-PCPI gave which is that the category 2 codes aren't really available outside of the PQRS program. The measures haven't been implemented in PQRS yet so rather than looking for the category 2 codes what we did was look to see whether the data elements are available in an EHR.

To give you a little bit more background about that testing there were two ways that the information was pulled from the EHR. The first was an automated report was pulled from the EHR and that really only looked to see whether data elements were available in structured standardized fields in the EHR. And then the second part of that testing was to do a manual review where someone went into the EHR and looked in other fields, not just structured fields. So they were able to go

into notes fields and look at attachments to see whether the information was available. And when you look at the testing information you'll see that we provide an automated rate and a manual rate.

After the measures were developed and tested and received full endorsement from NQF some of the measures were implemented in PQRS. And then this past year we pulled together another expert panel to review the measures again against current guidelines to update the measures and make sure that they are reflecting the current evidence.

We did -- the initial set that was originally endorsed I believe it had 12 measures. We did recommend to drop a couple of those because we thought that the information really isn't going to be captured in a standardized way across providers at this time.

And then we also combined two measures. We combined the chlamydia and

gonorrhea screening measure with the syphilis measure so that we have a broader STD screening measure.

I believe that's it. Thank you very much.

MEMBER MILSTONE: Thank you. So the first measure we'll be discussing this morning is 0412. This measure is titled "HIV/AIDS Hepatitis B Vaccination." This is not a new measure.

CHAIR SEPTIMUS: Aaron, can you get a little closer to the mike?

MEMBER MILSTONE: This is not a new measure. It was first introduced in 2008 as Jenna mentioned.

The measure assesses the percentage of patients aged 6 months and older with a diagnosis of HIV/AIDS who received at least one hepatitis B vaccination or who have documented immunity. This may sound familiar to a similar measure we discussed yesterday in the hep C population.

In this measure the numerator includes patients who have received at least one injection of hepatitis B vaccination or who have documented immunity. The denominator includes all patients aged 6 months and older with a diagnosis of HIV/AIDS with at least two visits in the measurement year with at least 90 days in between each visit.

This comes up, the 90 days between each visit issue is important because this is what drove the decision at least in the measure documentation to select or to choose one dose instead of three doses. There's concern because for hepatitis B vaccination there's a minimum amount of time required for the three-dose series where the first and the third dose have to happen at least 16 weeks apart. There has to be a 4-month window. And because of concerns that patients may drop out of care within 4 months it was decided that they would capture one dose to measure the start of the series in those with documented immunity --

I'm sorry, in those without documented immunity.

So in terms of -- that's the background. Should I move onto the impact?

CHAIR SEPTIMUS: Please, go ahead. We'll vote on -- just like we did yesterday. Impact, evidence and opportunity.

MEMBER MILSTONE: Sure. I think in terms of the impact I think in our work group there was consistent agreement that hepatitis B is a concern in HIV. I think there was consistent belief that hepatitis B vaccines should be given to all patients with HIV. I think the question from some of our members was whether or not the giving one dose of a vaccine is -- there's evidence to suggest that one dose of hepatitis vaccine will lead to the desired outcome. The desired outcome is immunity to hepatitis B in those with HIV and this is the same issue from yesterday with whether or not one dose will reach the intent of the outcome.

Maybe I'll leave it there for some discussion given what we -- since we had a long discussion about this yesterday.

CHAIR SEPTIMUS: Okay. So we're going to talk about the impact. So, Tom?

MEMBER FILE: Actually, and as our discussion yesterday about the validity of one dose, as I look at it we're not looking at it to see if one dose is an adequate immunogenetic, or provides immunogenicity for protection. We're just looking to see if that's a surrogate marker for if they're likely to receive all three versus never receiving one. I mean, that's the way I sort of look at it. And again, it goes to this measure burden that John talked about yesterday.

CHAIR SEPTIMUS: Mohamad?

MEMBER FAKIH: I fully agree with Tom. The only problem that I have is that we have to be consistent compared to yesterday.

Another thing that's important we never talked about is that hepatitis B and A

come as a combined vaccine also which, you know, we passed it for A so a lot of these people are going to get A and B at the same time. And we're not looking at, you know, identifying how many patients got B vaccine.

CHAIR SEPTIMUS: Michael, did you want to speak?

MEMBER FARBER: Yes. I just wanted to reiterate also that, you know, all, you know, for ACIP all high-risk patients are recommended to get more than one vaccine. So I guess I'm concerned if we make the recommendation in other words that the one vaccine would seem adequate. And that's to me how many people would interpret this so that I think that there should be I think at a minimum one vaccine would be useful but the committee should recommend just as all high-risk patients to get -- I don't see a reason not to continue vaccination considering the disease will go on for decades.

CHAIR SEPTIMUS: Okay. I think

the one vaccine issue is going to come back when we talk about reliability and validity. So, any other comments about impact? Obviously this is a population which is considered a high-risk population for hep B. I think Aaron's gone over that data. Is there any other comments relating to the impact of this measure?

(No response)

CHAIR SEPTIMUS: So I guess we're ready to vote. Okay. You remember how to use the clickers, right?

MS. KAHN: Voting on 1(a), high impact. It's high, moderate, low or insufficient evidence. Go ahead and start.

CHAIR SEPTIMUS: Jeff, are you able to vote online?

MEMBER BEAL: Yes, I think I am.

CHAIR SEPTIMUS: Great, thanks. I guess we're going to vote again.

MS. WINKLER: In the interest of time.

CHAIR SEPTIMUS: Everyone who believes it's a high impact raise their hands.

(A show of hands)

MS. WINKLER: Five high.

CHAIR SEPTIMUS: Next we'll go to moderate. Only vote once, guys.

(A show of hands)

MS. WINKLER: Eleven.

CHAIR SEPTIMUS: Low.

(A show of hands)

MS. WINKLER: One.

CHAIR SEPTIMUS: Insufficient.

(A show of hands)

MS. WINKLER: One insufficient.

CHAIR SEPTIMUS: Okay, so that passes. So let's go on to the evidence here.

MEMBER MILSTONE: Okay, so in terms of the evidence, again this was -- in discussion with our group was felt to support. I think that yesterday's discussions, one of the reviewers I should say said that this is based on the need for a hepatitis B vaccine, not

supporting one dose of vaccine as a measure. I think that was part of the questions that came up throughout is all the data presented really were based on hepatitis B vaccine as a prevention strategy for preventing hepatitis B in all patients, not just those with HIV. But there wasn't any direct data presented looking at the efficacy of one vaccine to prevent the outcome of hepatitis B.

CHAIR SEPTIMUS: Comments on the quality and quantity of the evidence in this population with one dose. I guess this group is ready to vote on the evidence. So is it working? You want to try it again? Okay. Now, remember this one -- you'll tell them how to do it in case they forgot. This one's a yes and no.

MS. KAHN: Okay, so voting on 18, evidence. It's yes, the body of evidence meets the guidance, no, the evidence does not meet the guidance, or no, insufficient information was submitted. So you can go ahead and start.

CHAIR SEPTIMUS: Okay, so this one -- it does not pass. I guess we have to record into the record. Okay. But do we need to read into the record the actual votes? Why don't you give us the votes?

MS. KAHN: So it's five yes, the body of evidence meets the guidance, five no, the evidence does not meet the guidance, and seven no, there's insufficient information submitted.

CHAIR SEPTIMUS: Okay, so remember -- of course this is one of the stop votes so that since it failed the question would be, as we did in a couple of them, do we want to make an exception for this particular measure. Aaron, you want to -- the developers. Which one do you want, Jenna or Bob? Bob and Jenna.

MS. WILLIAMS-BADER: We did have quite a lot of discussion about this with our expert panel and there were definitely experts who wanted to see all three for the reasons

mentioned. We did decide to go with the one dose because it did reduce measure burden and it also aligned with other measures that were also NQF-endorsed.

But I did want to comment that we are willing to take back to our experts a revision to the measure that would require all three doses rather than just the one.

CHAIR SEPTIMUS: Tom.

MEMBER GIORDANO: Could you please remind us how this -- measures with similar one-dose metric were handled yesterday?

MS. WINKLER: You did not pass it for that reason, for the hepatitis C population.

MEMBER GIORDANO: Okay. And there was no exemption granted.

CHAIR SEPTIMUS: I don't know if we had that discussion with that measure or not but it did not come up.

MEMBER GIORDANO: It did not come up. Okay, thank you.

MEMBER BEAL: May I make a comment, please?

CHAIR SEPTIMUS: Sure.

MEMBER BEAL: Okay, this is Jeff. I might suggest to the people making the measure that they consider perhaps changing the concept entirely to asking in the setting of HIV and AIDS for documentation in the medical file of a hepatitis B surface antibody quant or the hepatitis B surface antibody -- well, if they go after the quant then we really get what we want out of the vaccination is what I'm trying to say. I know there are screenings that talk about is antibody present or not before the vaccine, but if they eliminate the 1-2-3 vaccine and just go for the test of response of the vaccine they might get more meaningful data. Also not everybody responds. Just a thought.

CHAIR SEPTIMUS: Just to remind the group we did go online and routine antibodies after vaccination is not recommended for all

populations but it is for healthcare workers. This again -- but it might not be an unreasonable thing for this population as well. Aaron?

MEMBER MILSTONE: I just wanted to give a little more feedback as well, just some other comments that came up. I mean, I think there was also a clear discrepancy between the automated measure, automated validation and the manual validation. There was a 60 percent difference between what was found in EHR versus manual so I think that was a clear concern.

And then the other thing, and I'm going to bring this up later so I'll just introduce this concept now. You know, we spent time yesterday talking about these CPT codes and this measure relied heavily on the use of CPT codes for identifying patients with documented immunity. And I just think it would be important if you're considering revising the measure as to how you would either adapt that or find out how else you could capture

the information about whether or not patients had been immunized to hep C.

One of the comments that came up on the work group was if someone came into your practice 10 years ago with HIV and got a hep B vaccine series 10 years ago before EHR is it likely that in the current year they're going to document a CPT code for evidence of immunity to hepatitis B? Probably not. So we had concerns about the validity as well.

CHAIR SEPTIMUS: Peter?

MEMBER HAVENS: To that same end since there is a recommendation for universal hepatitis B vaccination for -- especially for younger kids now that age cohort is aging up into this population. And the history of vaccination is crucial to be able to opt out of this test and will not be easily captured. So that if the developers really want to test the adequacy of care in this regard they need to figure out how they will adequately capture either electronically or otherwise the stated

history of hepatitis B vaccination which will occur in a large cohort of younger patients many of whom will fit into this grouping and not need vaccination and potentially not have -- not reach the CPT or other criteria.

CHAIR SEPTIMUS: You want to say something? I want to wrap this up because I think that it sounds like we have some suggestions for the developers but that this, reconsidering this measure as an exception I'm gathering is not a strong opinion to do that. But Jenna, go ahead.

MS. WILLIAMS-BADER: Great, thank you. A couple of points. First, after some discussion with NQF yesterday we did want to let the committee know that we do have some form of e-specifications for these measures since they were tested in an EHR. And we would like to be able to provide those to the steering committee sometime soon in the future. They won't be available today obviously but since the measures were tested in an EHR and we have

those specifications if you'd like to consider them as EHR measures rather than the category 2 code measures then I think that option is on the table.

As far as the documenting whether or not a patient has been immunized, I think an important point is for the category 2 codes, first of all, you do have to report the category 2 code annually. That's in terms of participate in the program. That's what CMS is going to ask you to do.

But I think underlying that we do expect that a provider would know which patients have been vaccinated and which ones haven't. So you wouldn't necessarily have to document every year but you should review that yearly and make sure your patients are vaccinated. Otherwise if you don't you might not know which ones are vaccinated and which ones aren't.

I don't know if I want to get into the testing because if it gets to that point

we can address that.

CHAIR SEPTIMUS: Right. I'd like to just -- I want to wrap this up because I think this measure is going nowhere and I think we have some -- I think it's an important measure but it needs to be reworked and sent back to us when those changes are made. Does anybody really, I mean seriously need to make another comment about this measure? Because otherwise I'd like to move onto the next one since I think we've sort of beaten this to the ground.

Anybody else? Seriously, I don't want to cut off discussion but is there anything we haven't said that needs to be said? Well okay, we thank you.

Let's go onto the next measure which is 0404, HIV/AIDS. I think Kathleen, do the developers have anything in addition they want to say about this measure or just let Kathleen discuss it? Jenna? We'll be nice to you.

MS. WILLIAMS-BADER: No, I don't

think we have anything additional to say.

CHAIR SEPTIMUS: Kathleen, I think this is yours.

MEMBER BRADY: It is. Okay, so the title of this measure is "HIV/AIDS CD4 Cell Count or Percentage Performed." The brief description of the measure is percentage of patients aged 6 months and older with a diagnosis of HIV/AIDS with a CD4 cell count or percentage performed at least once every 6 months. The numerator is patients with a CD4 cell count or percentage performed at least once every 6 months. And the denominator is all patients aged 6 months and older with a diagnosis of HIV/AIDS who had at least two medical visits during the measurement year with at least 90 days between each visit.

In terms of impact, I mean there's about 1.2 million people in the U.S. living with HIV and AIDS. And monitoring CD4 cell count in HIV is one of the key factors in deciding -- you know, actually not really

anymore -- when to initiate antiretroviral therapy. But it has been in the past, but certainly for prophylaxis for opportunistic infections. It's a strong predictor of disease progression and survival.

So, and I don't think -- and I think for the most part in our work group for the most part everyone thought the impact was either high or moderate.

CHAIR SEPTIMUS: Okay. Any comments on impact? If not we'll go to vote on impact.

MS. KAHN: Voting on 1(a) high impact, high, moderate, low, or insufficient evidence. You can go ahead and start. You have 13 high, 4 moderate, 1 low and zero insufficient evidence.

CHAIR SEPTIMUS: Okay, let's then go to evidence.

MEMBER BRADY: Okay, so for evidence -- it might help if I was actually on the right measure. Okay, for evidence, you

know, it's for the most part most of the studies are not randomized controlled trials but cohort studies. There were seven studies cited in the current DHHS guidelines. Five were cohort studies of 16,446 patients and 2 were control studies, case-controlled studies including 48 patients. So, I mean there's a fairly large amount of evidence regarding this.

CHAIR SEPTIMUS: Comments on the evidence? So you found a fair number of studies but there wasn't a randomized controlled trial.

MEMBER BRADY: There's not, no. For the most part it's based on cohort studies.

MEMBER HAVENS: In pediatrics there are randomized controlled trials suggesting that monitoring frequency can lead to differential implementation of antiretroviral therapy. So, for children the level of evidence would be high. Unusually for adults the level of evidence is less.

CHAIR SEPTIMUS: Okay. Again, for

this one it's going to be yes, there's evidence, no, there isn't, or three, it's insufficient. So, but I was just going through the quantity and quality of the criteria that NQF uses. Any other comments before we vote on the evidence? Okay, we'll vote on the evidence.

MS. KAHN: Voting on 18, evidence. You can go ahead and start. You have 15 yes, the body of evidence meets the guidance and 3 no, evidence does not meet the guidance, and zero no, insufficient information.

CHAIR SEPTIMUS: Thank you. Now we go to opportunity and gap.

MEMBER BRADY: Okay, so the data submitted for performance gap was from the 2009 and -10 CMS PQRS system for which the average performance rate per eligible professional was 76.8 percent in 2009 and 83.9 percent in 2010. And developers report they feel that is an indication that there's a gap in care with room for improvement.

I will note that the measure is not

stratified by patient groups or cohorts and that our work group felt that that was something that was lacking.

CHAIR SEPTIMUS: So nothing about disparities in this group at all?

MEMBER BRADY: No.

CHAIR SEPTIMUS: Okay.

MS. WINKLER: In general do we know that this is a particular area of disparities in care?

MEMBER BRADY: Yes, we know that from data. There's actually CDC has released data as well as actually I was someone who participated in a four-city analysis from HIV surveillance data in the Medical Monitoring Project that indicated that there were significant racial and ethnic disparities in HIV treatment.

CHAIR SEPTIMUS: Yes, Peter.

MEMBER HAVENS: So this measure requires two medical visits during the measurement year. One of the problems that

is identified is that -- that we'll get to in some of the other measures is the fact that people don't come back. They're seen once and don't come back. So they don't either get a repeat visit or a CD4.

So it could be argued that the apparently high percentage of testing here overestimates the true activity and when you look at this measure in combination with the visit frequency measures that we'll be reviewing later, that you might actually get a more complete picture of the inadequacy of care delivered in many different populations. So that while 85 percent compliance with this testing frequency may look good, when you combine this with the other information on visit frequency this is already a group of people who are coming back. So.

MEMBER BRADY: Yes, but I think actually, you know, the way the recommendations are is that persons who are stable can get a CD4 and therefore also viral load measurement

every 6 to 12 months. But you know, this doesn't break that out. So you have lots of people who probably have detectable viral loads who may be only coming in and may have low CD4 counts that are only coming in, you know, once a year. And it's going to look by this measure that they, you know, they're meeting it. Or they may come in twice a year. And they look like they're meeting the measure even though they're not getting adequate care.

CHAIR SEPTIMUS: Any other comments? Jeff, I know that you were in this work group so if you'd like to make a comment just speak up, please. Tom?

MEMBER GIORDANO: So, to follow up on that comment I guess I would -- I appreciate the fact that this is among people who are engaged in care at least at a minimum level by having two visits each year. I think it allows the organization that's using it to hone down a little bit on the actual measure which is did they get a CD4 count done if they were

in care enough.

There are other measures that get at whether people have enough visits that we'll look at later but I think this is -- personally I think this is the right denominator. If you broaden it to everyone who had any visit in the year then you get a mixed bag of performance that you're measuring.

CHAIR SEPTIMUS: Kathleen?

MEMBER BRADY: Actually, I don't really have a problem with the denominator. I have a problem with the numerator which is that it's at least every 6 months. You could have somebody who comes in in January and in June who's stable but on therapy undetectable for 15 years. They're not going to meet this measure because it's at least every 6 months. They're not going to meet the measure.

No, the visit. They come in, they meet the -- they end up in the denominator. They will not end up in the numerator. If they come in July or -- but then you could have

somebody who meets the measure who comes in in January and then in December and they're essentially getting care once a year.

CHAIR SEPTIMUS: I think we're going to get into -- as we get into the reliability and validity.

MEMBER BRADY: But that's a problem I have with the measure.

CHAIR SEPTIMUS: Why don't we go ahead and just maybe focus right now on -- does anybody have any other comments relating to the performance gap? We can vote on that and then we can get into reliability and validity. So are we ready to vote on the gap? Let's vote on the gap.

MS. KAHN: Voting on 1(b), performance gap. You can go ahead and start. We have 2 high, 16 moderate, zero low and zero insufficient evidence.

CHAIR SEPTIMUS: Okay, Kathleen. Now we're going to talk about the reliability and validity which I think gets into some of

the comments that were just recently made.

MEMBER BRADY: Okay. All right. So the first comment that I'm going to have about this if I get to the right section is that in terms of the numerator details it says that it's patients with a CD4 cell count or percentage performed at least once every 6 months. This came up at the -- during our call that it's either a CPT procedure code or report of a CPT category 2 code that it was documented which I think means that it was just ordered, is that correct?

MS. WILLIAMS-BADER: I believe -- I'm just looking to the category 2 code right now. Sorry.

MEMBER BRADY: Because it says CD4 cell count or CD4 cell percentage documented as performed, but doesn't that just mean there was an order placed for that?

MS. WILLIAMS-BADER: Performed means that it was actually completed, that you know that it was done. Otherwise we would have

said ordered if it was ordered.

MEMBER BRADY: All right. So certainly one of the things that's come up about this is that once again it's using potentially the CPT category 2 codes which are infrequently reported. And so I think the other issues were the fact that there's this difference between the numerator and denominator in terms of who gets in.

And in terms of reliability. So, I'm just trying to scroll down. Sorry. All right, denominator details, so yes, it's using ICD-9 codes for the denominator. So, I don't know, it seems -- the denominator details seem somewhat complicated.

CHAIR SEPTIMUS: Can I ask, this is a maintenance measure, correct?

MEMBER BRADY: Yes.

CHAIR SEPTIMUS: So what has been your experience with measurement of this? Or do you have any?

MS. WILLIAMS-BADER: I believe

this is one of the measures that is included in the PQRS program. So, we do -- I mean there are providers out there who are reporting the measure using it as specified.

CHAIR SEPTIMUS: Aaron?

MEMBER MILSTONE: Thank you. So my questions I alluded to earlier have to do with the use of CPT codes. These were data that were validated in four sites in the Midwest region when it was originally done. And I'm curious if you have data on how reliable and valid the use of CPT procedure in the CPT category 2 code as reported are in detecting patients that actually have this done.

MS. WILLIAMS-BADER: No, we have not actually tested the category 2 codes themselves. As we said in using the process, the protocol that AMA described yesterday these were tested in an EHR rather than testing the category 2 codes themselves.

MS. BURSTIN: So, just to clarify it's the same issue as yesterday. You

essentially only have the EHR-based testing and so only the e-specs would actually be endorsed at this point.

MEMBER MILSTONE: But this one actually doesn't list anywhere in here as an e-measure. I mean even if you look at the numerator it doesn't even have --

MS. BURSTIN: Correct. And that will be adjusted.

MEMBER BRADY: There's no e-specifications.

MS. BURSTIN: Right and those will be submitted to us from PCPI. This was a joint measure they did jointly.

MEMBER BRADY: So there is information presented that results comparing electronic health record automated report to visual inspection of the medical record. And the automated calculation of performance was 80.5 percent whereas manual calculation of performance was 90 percent for a difference of 9 percent.

CHAIR SEPTIMUS: Actually for some of the things we're going to be discussing that's not too bad.

MS. BURSTIN: And those measures initially came in as time-limited meaning they didn't have testing at the time. Those testing results were submitted, reviewed by our Consensus Standards Approval Committee and approved.

CHAIR SEPTIMUS: Aaron?

MEMBER MILSTONE: So I guess my question then is how do we know what the -- so what were the e-measures? In what group was that assessed in for reliability and validity? Was that -- no I know the data but what was the size of the -- what was the population? Is it in there?

MEMBER BRADY: It was 1,465 patient encounters. And it was in the Midwest and it was performed in 2009. And it was four sites representing community health centers serving primarily low-income and uninsured patients.

MEMBER MILSTONE: So just to clarify that's one small population using one electronic health record. So it doesn't -- we don't have data on how this performs using the e-specs in other electronic health records.

CHAIR SEPTIMUS: Michael?

MEMBER FARBER: I wanted to just make a comment that from my Medicaid experience that we don't talk about ordering, we don't talk about billing, we only talk about what's reimbursed. So once it's paid then it's assumed it's done.

With ICD-9 that's where it's a problem because usually there's no money attached to it. So, this is what I saw yesterday a lot of issues. When you talk about an ICD-9 code you don't really have proof in any way that they have the diagnosis that's specified unless you do some internal review. But as far as the, you know, a CD4 count you would want to see it reimbursed. Then you would assume it's been done.

MS. BURSTIN: If I could just respond to Aaron's initial -- the last question and it's a good one. We're actually I think very much still in the process of trying to understand what testing is required for EHRs. At this point we have not required more than one EHR system to be evaluated. Partly because I think you've seen one, you've seen one. It's not clear how many you need to actually understand this. So I think as we're getting more experience with it we'll have a better sense of how to proceed. But this is active work that the Office of National Coordinator is doing, that others are doing, of trying to figure out exactly how to do it. But based on what was provided it met our bar. We'd certainly like it to be higher, we'd all like it to be higher and I think as we get a better sense of the best way to test those measures I think we'll have a better understanding of how to proceed.

CHAIR SEPTIMUS: Yes, Tiffany.

MEMBER OSBORN: I have to say I feel very uncomfortable with that. I don't feel comfortable with that at all because physicians and hospitals are going to be -- their reimbursement is impacted specifically by how these are measured. And I think what Aaron has brought up is critically important. And that we have to look at whether or not these measures have been presented to us as reliable and valid in measurement for what we have in front of us. And to -- I'm not -- so, do you --

MEMBER MILSTONE: No, I appreciate your comment. I mean, I'm struggling because I have been talking over the last day or two with primary care physicians about this who say, you know, I have 15 minutes to see a patient or 10 minutes and if I don't check that box or the CPT code that I don't even know exists I'm going to get not reimbursed for an aspect of my visit.

And so I think as a clinician I think

what we're doing right here is very important. And I understand the importance for quality improvement, I understand the importance for benchmarking and reporting, but I think from the other side is I want to make sure that what we're saying is acceptable that will impact the livelihood of clinicians is in fact reliable and valid.

MS. BURSTIN: And I think it's why you're only looking at the EHR testing which was done. We don't have CPT-2, exactly to your point. We don't know the reliability and validity of the CPT-2 based collection which is why at this point we're only looking at the EHR testing that was provided.

Again, you need to vote your conscience but I, you know, at least as part of what our testing task force put forward what was submitted was adequate. That's all I can say.

CHAIR SEPTIMUS: Okay. Peter and then back to Tiffany. But remember we're --

we want to discuss reliability first and then we're going to talk about validity. So I know these things overlap tremendously but we do need to make sure we have to go in that order. So, Peter?

MEMBER HAVENS: I was just going to reaffirm the need for NQF to make a strong pitch to anybody who brings these. And you can lead in this regard by demanding more testing across more EHRs and require, recognizing that reliability and validity are crucial markers for further review of existing measures.

CHAIR SEPTIMUS: Well said. Tiffany?

MEMBER OSBORN: Just to sort of go back to what we talked about yesterday. I mean, we had a very long discussion regarding severe sepsis and septic shock. And the reason that that did not pass was not because of the scientific validity or the scientific evidence, it was because of reliability and

validity in the measurement process. And I think that that needs to hold firm.

And really quite frankly I don't think most of us as clinicians mind, we all want quality. But I don't think we mind getting docked for something we didn't do well. What we do mind is getting docked for something that wasn't measured well, or defined well, you know. That's really problematic.

MS. BURSTIN: And we agree with you completely. And I just think we need to be consistent. The measure testing you looked at yesterday for hepatitis C was done in a very similar process and at least for the measures you put forward you deemed them acceptable. Really, just to be consistent from day one to day two, in addition to the fact that we're actually going to return to the sepsis measure with some additional discussion later today.

CHAIR SEPTIMUS: Okay, Tom.

MEMBER FILE: Just very quickly along the lines of what Tiffany and Aaron have

brought up I think which is very important. I just want to clarify. We can get input from our NQF colleagues here. When we look at the total endorsement process, I mean we are a steering committee. I mean, what we say is not the final answer obviously. It'll go through public and member comment and I assume even the developers can come back and make comments and changes or whatever. And I think these types of issues are extremely important. And to what extent are these evaluated by the potential users in this whole endorsement process. And can I ask you what percentage of the sort of measures that we approve actually are significantly changed when you get to the bottom line for definite completion of the measure?

MS. WINKLER: That's also an evolving issue, the number of measures that have changed. In the early years when measures were less well-formed and well-constructed there were often a lot of malleability to them.

However, now that the requirement for testing is as solid as it is right now if you start changing the measure your testing does not apply. So it's become less of an issue and that's why we're asking you to really evaluate what you have in front of us.

MS. BURSTIN: Absolutely. And also as the measures get out and they're in use and there's implementation and experience, and we learn where there are issues, again, just like I said yesterday, if there's a change in evidence we'll do an ad hoc review. We'll also do an ad hoc review anytime there's evidence of implementation issues in the field as well. This isn't actually adequately measuring. The developer makes a material change to the measure. We'll re-review it again. And again, I think as a lot of these measures are being put out, meaningful use, other issues, HRSA programs as we'll hear in a bit, I think we'll get much more experience in how they perform.

CHAIR SEPTIMUS: Okay. So I think, unless -- I don't see any -- okay. Let's again first vote on reliability. As you know the -- are there precise specifications and evidence of reliability either in data elements or measured score. So that's the first thing. Then we'll go to validity if this measure passes.

MS. KAHN: Voting on 2(a), reliability. You can go ahead and start. We have zero high, 11 moderate, 4 low, and 4 insufficient evidence.

CHAIR SEPTIMUS: Okay, so this passes. Now we're going to go to validity where I think a lot of the comments may -- Kathleen?

MEMBER BRADY: Yes. So this goes back to, you know, are we measuring really what we want to measure and is the data or the measure consistent with the evidence. And I would have to say, you know, based on the numerator issues that I mentioned earlier that I don't think

that it is, that you can -- there could be significant misclassification.

MEMBER SPACH: Can you elaborate?

MEMBER BRADY: Can I elaborate? So that would be the example of someone who comes in in a January and June, you know, because it has to be every 6 months. So you know, even if you're off a few days you're going to be put into a category that you didn't meet the measure, you know. And then someone who is essentially seen only really once a year, that person who's seen in January and then December is then actually included in the numerator as meeting the measure when they've only been really seen once a year. I think it's just that the numerator definition is sort of too tight and it should be maybe more of a range. And so.

CHAIR SEPTIMUS: Tom?

MEMBER GIORDANO: I appreciate that comment. I would add though that it's -- having thought about how to measure

retention in care which is sort of analogous measurement issues to CD4s, it's incredibly difficult to come up with a measure that is -- it's impossible to come up with a measure that's perfect.

You're always going to misclassify some people. In the denominator they have the 90-day rule to try to make sure that the visits are spread out a little bit but it's true that you could misclassify. Even if you adopted that rule for the numerator you could misclassify someone who had a visit in January and December. They would be considered meeting the measure when in fact it's not optimal care probably.

But so every 6 months is, I don't think it's adequately defined in the measure as it's presented. I don't know exactly what that means when you operationalize it. Does it mean that you have to have a visit at least 180 days or at least 6 months from the first one or exactly in or at some window around the

180-day anniversary? I don't know how to operationalize that based on what's presented. But you have to accept some error essentially because these things are very difficult to operationalize.

MEMBER BEAL: This is Jeff. I'd just like to comment that this has been discussed a great deal in HIVQUAL and I believe this is the definition that is used in HIVQUAL and this is also a HRSA performance measure. And this is the definition directly from the HRSA performance measures that we do for Ryan White program quality improvement.

And also, just to note for our group when we looked at this as validity as a group on a smaller conference call the majority of us felt that it was moderate in validity.

CHAIR SEPTIMUS: Let me ask the question slightly differently and then Jenna can respond. Does this measure -- can it be used to create for the physician or the clinic does it provide them with information where

they can see opportunities for improvement?

MS. WILLIAMS-BADER: So we did during the work group calls get this question about what exactly once every 6 months mean. And we would like to take that back to our expert panel and to clarify that definition because we realize that that is up to interpretation. And so I do think we would like to take that back to our experts and give them an opportunity to clarify what exactly that does mean.

MEMBER SPACH: This is David. I would just suggest that if you are going to take it back to the expert panel that you add in information that's consistent with the most recent DHHS guidelines regarding stable patients on antiretroviral therapy who have suppressed. And I can give you the exact, but the wording is something to the effect in such patients CD4 may be monitored every 6 to 12 months unless there are changes in the patient's clinical status. And that's talking

about patients who are suppressed on antiretroviral therapy. So, if you're taking it back that slight amendment would be a good exclusion so providers who have stable long-term patients wouldn't be penalized for this.

CHAIR SEPTIMUS: Peter?

MEMBER HAVENS: Just to point out that if we're going to make a change here that the -- what I was trying to point out before is that this is part of the suite of measures that look at retention in care and that while there are problems with all of them they allow, taken together, a broad view of the pattern of care. And this specifically look at one physician-related act that should happen when people are kept in care as Tom points out. And so that while the denominator is open to question in certain instances taken broadly I think it gives the best picture possible.

One approach would be to, as we move forward to get -- well, I don't know if this

is an approach that's possible, to stress to NQF that testing to see, to get more information on the validity of this measurement would be particularly important given the concerns of this group.

CHAIR SEPTIMUS: Okay. I see no others. I think we'll -- Aaron.

MEMBER MILSTONE: Just for some guidance from our chairs. So if there's questions about the definition in terms of going back to the committee and revising, what are we voting on then? How do we vote if there's questions about changes?

CHAIR SEPTIMUS: I'm going to have to kick this to a higher level on the food chain here. Karen or Reva?

MS. BURSTIN: I'm not God.

(Laughter)

MS. BURSTIN: So I was actually just asking Jenna as a sidebar how soon they could actually bring these questions back and it sounds like it's just a couple of weeks.

So I think this might be something appropriate to defer and come back for further discussion after they've had a chance to discuss with their committee.

CHAIR SEPTIMUS: So an option is that we stop here and reconvene the call after the measure has been reworked?

MS. BURSTIN: Yes, we will have to be quick about it. This is supposed to be out for comment in mid-September as Reva reminds me. She has to stick to the time lines. I get to play God. So, we'll have to make it quick. We'll have an offline conversation with NCQA. I mean, it's very targeted, specific questions and we'd come back to you in email to finish the discussion.

CHAIR SEPTIMUS: Okay. So, Tom.

MEMBER GIORDANO: If that happens won't we be in the position where you'll have a modified definition and we'll have no data validating that definition?

MS. BURSTIN: That's a concern but

I guess one question might be can you invoke if you can find it every 6 months. I mean I think the issue is more so in terms of the reliability of what you're looking at. In terms of timing I'm not sure the timing variable changes by changing the time period. We'll have to see what they bring back.

CHAIR SEPTIMUS: Okay. So, go ahead.

MEMBER MURTHY: I'm wondering if I'm just reading this correctly. I just want to clarify. If I'm reading this correctly there is a difference between the manual calculation from the automated calculation performance of about 9 percent. Is there -- do you have information whether that 9 percent difference is attributed to this kind of finding with the 6-month difference? Could that be an answer for us?

MS. WILLIAMS-BADER: There were two main reasons that were listed for -- that were provided as reasons for the gap. One was

that CD4/CD8 ratio code had made its way into the codes that were tested. That caused some confusion at the test site. That code has been removed because the ratio is not appropriate for this measure.

The other was the timing and what exactly was meant by every 6 months. So I think if we provide a clearer definition that would help with the reliability and validity of the measure.

MEMBER MURTHY: I'm sorry, and what about the performance gap itself? The 90 percent versus 100 percent. Is there a sense for how much of that may be impacted by the timing definition?

MS. WILLIAMS-BADER: We didn't look at that, no.

CHAIR SEPTIMUS: Compared to some of the measures we're going to talk about 9 percent is pretty good. So I'm not -- okay.

So we have two options here. One is to stop here, let them modify things based

on our conversation or go on and vote on the validity measure and let it go the way it goes. So, I guess by a show of hands who would like to stop here and wait for them to revise this and then take this back up on a conference call? So raise your hands if you want to do that.

(A show of hands)

CHAIR SEPTIMUS: Sounds like you want to vote. Did I get that? All of you who want to vote now raise your hands.

(A show of hands)

CHAIR SEPTIMUS: Okay. So.

MS. BURSTIN: And even if you vote they can still bring back information and you can re-vote. So it's not a done deal.

CHAIR SEPTIMUS: Okay. All right. So let's go ahead and vote on validity then.

MS. KAHN: Okay, voting on 2(b), validity. You can go ahead and start. I think we're missing one person. If we could have everyone press it one more time. We were doing so well. Okay, there we go. Zero high, 10

moderate, 4 low and 5 insufficient evidence.

CHAIR SEPTIMUS: Okay. It slithered by. Okay. So let's keep going now. We've got usability and feasibility. So Kathleen, you want to take us by the usability?

MEMBER BRADY: Okay. So, the measure was used in the CMS PQRS program in 2009, -10 and -11. And that's really where it's been used. They do report that HRSA uses a similar measure but it is actually somewhat different. The numerator is different. And so that's, you know, something that they mentioned. And that's really all I have to say.

CHAIR SEPTIMUS: Any comments now on usability? Meaningful, understandable, can be used for public reporting. Okay, let's vote on this measure then.

MS. KAHN: Voting on usability. You can go ahead and start. We have 4 high, 10 moderate, 1 low and 4 insufficient information.

CHAIR SEPTIMUS: Okay. Next one is feasibility. One of the things about feasibility just to remind the group about inaccuracies and unintended consequences are in this particular element. Kathleen, any additional comments?

MEMBER BRADY: I don't really think that I have any other comments, you know, other than what we've talked about previously.

CHAIR SEPTIMUS: Seeing no comments we'll go ahead and vote on this element.

MS. KAHN: Voting on feasibility. You can go ahead and start. We're missing one person in the room. Can everyone press it one more time? Sorry. All right. We have 2 high, 11 moderate, 2 low and 4 insufficient information.

CHAIR SEPTIMUS: Excellent. Now we're going to read the last thing. Is the measure suitable for endorsement?

MS. KAHN: Does the measure meet

NQF criteria for endorsement? You can go ahead and start. You have 11 yes and 8 no.

CHAIR SEPTIMUS: Thank you very much. I want to -- I really like what Peter said earlier in that I think when we start talking about some of the other measures about visits, et cetera, when you build that whole number of elements together I think it gives you a very true picture about care that's being provided. So we need to keep our mind on that.

Is Diane Jacobsen on the call, Operator?

OPERATOR: Diane Jacobsen is on the call.

CHAIR SEPTIMUS: Good morning, Diane. It's Ed.

MS. JACOBSEN: Good morning, Ed, how are you?

CHAIR SEPTIMUS: Fine. Okay, so we're going to go back to 0298, "Central Line Bundle Compliance." The developer is IHI. So Diane, if you'd make a few comments and then

Mohamad is going to discuss this. Diane?

MS. JACOBSEN: Thank you very much. This conversation has been incredibly helpful. I appreciated having the opportunity to be part of it, particularly the discussion related to the sepsis bundles yesterday which are measures I'm very, very familiar with also. I think the challenge with the central line bundle is around the reliability and validity in the measurement process which has been discussed a great deal. And the intent of this measure developed as a reliability measure. It was not intended to address or include all the elements of care related to the central line but rather a small group in a bundle that when taken together promote teamwork, collaboration and other influences that ultimately have been shown to affect the outcome measure of central line-associated bloodstream infections. So I wanted to just state that.

I really appreciated the comments

that were included in the preliminary evaluations and agree with them. That said, many hospitals, many systems have utilized the bundle measurements as a process measure which is how they were developed and intended, and clearly have been useful in facilitating improvement across organizations. So with that I appreciate any discussion and feedback and will respond to any questions.

MEMBER FAKIH: Thank you very much, Diane. This is Mohamad Fakih. I think whatever you mentioned I fully agree with. I think the impact of every single item in the bundle, not every single item but most of them is, you know, they're very important. So the chlorhexidine use, the complete barrier, all of these are supported by IDSA.

You know, so as far as an impact of the individual points that are part of the bundle, you know, I think they're high-impact and all of them are between 1, almost all of them are category 1 as far as evidence.

The issue is that this is, the tool is -- you know, what's asked is the documentation. The tool itself is just documentation that these were done. And one of the questions that I've had is how accurate is the documentation. And this is something that we cannot, you know, I didn't see any literature about the accuracy of documentation of that tool. So whether it reflects really what happens at the bedside when the operation is done.

MS. JACOBSEN: May I comment on that?

MEMBER FAKIH: Absolutely.

MS. JACOBSEN: I agree with you and one of the things in the, you know, submitting this. I did reach out. There are currently two states that include the central line bundle as one of their publicly reported measures. One of those states happens to be Minnesota which is where I reside. And I spoke with them and they raised that question also. The data

is self-reported by the hospitals and the feasibility of validation or reliability hasn't been feasible in those states. It's dependent upon the individuals within the hospitals collecting the data. So I think this is an important point for consideration with this type of measure involving central line bundle obviously but also as discussed yesterday sepsis and the ventilator bundle which we did withdraw for consideration. So it really is a challenge.

CHAIR BROTMAN: Okay, thank you. And let me speak to impact as well. So why don't you start by going through the impact, Mohamad, and we'll go through systematically.

MEMBER FAKIH: So again, you know, the impact. I mean, it's multiple parts. One of them is the complete -- it's a bundle. I can read the bundle for them and what? Okay. So hand hygiene, maximal barrier, precautions upon insertion, chlorhexidine skin antisepsis, optimal catheter size selection with avoidance

of femoral line. There were a couple of articles, one of them is a meta analysis that shows, that reviewed femoral versus IJ, you know, internal jugular, and did not see much of a difference. The, you know, the IHI bundles states that avoid the femoral line. A lot of articles in the past have, you know, recommended an effect idea, say -- also has recommended not using the femoral line as, you know, as central line because of a higher risk of infection.

Daily review of line necessity. That's another part of the bundle that's tough to measure, the daily review of line necessity. How to document it. It's a great thing to do, we should do that, but it's very tough to obtain that data element.

CHAIR BROTMAN: Let's just stick with the impact of this process, of looking at this.

MEMBER FAKIH: So the impact is very high as far as certain components such

as chlorhexidine antisepsis. You know, if you use it versus betadine it's much better as an antiseptic agent and decreases the risk of central line infection.

CHAIR BROTMAN: But the impact overall of having a bundled package to address this potentially extremely serious situation?

MEMBER FAKIH: Okay. So you know, if we look at the whole bundle right now I think the impact is probably low to moderate. Just let me -- so I'll explain the reason why. Because right now we have other measures that are in place that give feedback to hospitals. So using that bundle, I'm talking about that sheet, not the steps, that sheet, I don't think it has a huge impact at least that I can see.

CHAIR BROTMAN: Anybody want to comment on Mohamad's? Tiffany.

MEMBER OSBORN: Regarding impact I think that most studies where they have implemented this they've seen a fairly significant improvement in central venous

catheter infections. So, and that is -- you're talking, what, an estimated $34,000 -- I mean, there's a significant impact both in cost, in lives.

And all of the studies that I've seen to date, I might have missed some, but all the studies I've seen to date that have implemented this bundle have found both a survival benefit and a cost benefit. So, I mean we can argue about other components of you know, of the bundle but as far as the potential for impact I think that the potential for impact is quite high.

MEMBER FAKIH: Just to clarify I am not debating. The bundle itself is excellent. It's the documentation, using documentation of bundle. And this is -- so there are two different issues in this case.

MEMBER OSBORN: But right now we're just talking about impact.

MEMBER FAKIH: Okay, impact is high. Impact is high. Impact is high.

CHAIR BROTMAN: Right, I just wanted to refocus you on that. I don't know, Aaron, if you want to talk about the point of those studies and so forth regarding the impact.

MEMBER MILSTONE: Sure. I was just going to reiterate what Tiffany said which I think the evidence is clear that -- including the Peter Pronovost New England Journal study that was done in the Keystone collaborative in Michigan. I think there's little question in the field of healthcare infection control that the bundle has been a dramatic driver of reductions in infections. I think you're getting at whether it's measuring the bundle versus the bundle itself but I think the impact is clear.

CHAIR BROTMAN: I just want to make sure we isolate the impact because the impact I think is extremely clear for a lot of us. If there's no other discussion let's go to vote. I'm sorry, Adam? Okay.

MEMBER THOMPSON: Yes, and I was just going to say from a patient viewpoint on this this is something we just went through with my mother and it's an easy thing that you can give patients to check up on the care of not only when they're getting one but also on their family members because that checklist is something we monitored with my mother very carefully when she was in it. And so it's a tool that patients can use as well.

CHAIR BROTMAN: I agree with that. Thank you. All right, let's go to vote on high impact.

MS. KAHN: Voting on 1(a), high impact. You can go ahead and start. Eighteen high, one moderate, zero low and zero insufficient evidence.

CHAIR BROTMAN: Okay, so that passes. Let's go to the evidence. Mohamad?

MEMBER FAKIH: So again I had mentioned like the chlorhexidine antisepsis is much better than betadine complete barrier.

And you know, Pronovost study, you know, this is the Keystone study. It had also another element which is cost. So the teamwork, you know, I think Diane has mentioned that is another part. But the evidence is also high that it does work.

CHAIR BROTMAN: The evidence presented in the specifications?

MEMBER FAKIH: I mean, this is again category 1. I can tell you with the IDSA recommendations a lot of the stuff that they are mentioning are category 1(a) or 1(b). So avoiding femoral line is 1(a) from IDSA. Aseptic technique, you know, maintaining septic technique is 1(b). So, all of this, all of those are high evidence. There are a few that, I think the data evaluation is not but many of those -- the chlorhexidine is a 1(a) category from IDSA guidelines.

CHAIR BROTMAN: I think I murkily recall there was one discussion point about the checklist is a great tool, but changing

the culture in the hospital is also extremely important. And to that point that having the checklist may actually add to changing the culture within the hospital, seeing the improvements and so forth. And I think it's been that way for a number of institutions.

MEMBER FAKIH: You know, but most of these studies were done with cost implementation. There are other hospitals that have used other high-reliability tools such as, you know, maybe another high-reliability tool other than CUSP. I am not -- I mean, I can't tell you if it's, if the tool itself really changes the behavior because it was always compounded with something else with it.

CHAIR BROTMAN: Right. So let's just stick with the evidence. Was there any other discussion? Ed.

CHAIR SEPTIMUS: This is not necessarily against the bundle but I just want to raise an issue about patient safety. It

has to do that most CLABSIs occur outside the ICU and that in fact the maintenance of lines may in fact be more critical than actually the insertion of those lines. That is not to say that the insertion and using an alcohol chlorhexidine prep is not important but I want to let you know that this is a small part of HAI prevention and most of these studies have been done in the intensive care unit. So I just, just a caution. I'm not against it but I want to let you know that this in and of itself is not going to get us where we want to go in patient safety.

MEMBER FAKIH: You know, there's a huge change in the epidemiology of central line infection. We used to have about 4 or 5 per 1,000 catheter dates, you know, as infection and now it's less than 1. And a big part of it is related to the insertion and now the main part becomes the maintenance because we're doing so good at insertion. So, but you know, I can understand that the developer said

that it's not to cover everything, the bundle. And again, I look at specific parts of the bundle, they're okay, it's just the documentation of the bundle is what I have a problem with.

CHAIR BROTMAN: The concepts for insertion and maintenance tend to overlap or they would, you know, sort of parlay onto each other depending upon I think what the future evidence shows. But a lot of times the lines are maintained the same way that they were almost inserted and that's been my experience especially at the home care level.

Any other discussion? Let's go vote for -- at the evidence point at this point.

MS. KAHN: Voting on 18 evidence. You can go ahead and start. Seventeen for yes, the body of evidence meets the guidance, two for no, the evidence does not meet the guidance, and zero no, insufficient information was submitted.

CHAIR BROTMAN: Okay, so that

passes. Now we need to get into the performance gap. Mohamad?

MEMBER FAKIH: You know, again, there's a huge improvement compared to before. I don't think I have that information about how much of a difference there is right now as far as the compliance with the bundle. I don't think I have that.

CHAIR BROTMAN: The measure developer didn't supply it. No data.

MEMBER FAKIH: Yes.

CHAIR BROTMAN: No data. All right. Was there any discussion in the work group that you remember specifically?

MEMBER FAKIH: I think we asked -- Diane?

MS. JACOBSEN: Yes.

MEMBER FAKIH: Do you have data about how much of a gap as far as compliance with the bundle?

MS. JACOBSEN: Well again, this is relatively, you know, a challenging question

in that these are self-reported measures. Many hospitals within collaboratives have reported their reliability and achieved high reliability with the overall bundle. But as far as public reporting, like I said two of the states, Rhode Island and Minnesota use this currently and they utilize self-reported data from the individual hospitals.

CHAIR BROTMAN: Okay, thank you.

MS. JACOBSEN: Does that address the question?

MEMBER HAVENS: No. The question is when people put in central lines in ICUs what percentage of people who use this performance measure report putting in those lines using a bundle. If that percentage is 50 percent then there's a big chance for improvement. If that percentage using a bundle is reported as 95 percent there's little chance for improvement. What is the current rate of bundle use in ICU patients? That's the question.

MEMBER FAKIH: You know, I can give you like, just an example. My hospital has been using the bundle as part of the Keystone. In 2003 we started doing this. Right now when I look at the sheets all of them are yes, yes, yes. None of them is yes with correction. So compliance is 100 percent and no mistake. And this is one of the issues that I have with this. But this is one hospital.

And I don't know what, you know, what you've seen in Minnesota or in these two states that you reported, how much is the compliance. Is it 100 percent? And do you think that, you know, with what Peter is saying do you think they're reporting on every single line? And that's another issue is reporting. Do you get all these lines inserted reported on in the ICU?

CHAIR BROTMAN: Aaron.

MEMBER MILSTONE: Yes, I think --

MS. JACOBSEN: Clearly organizations that have utilized the bundle

state the compliance is very high. And a great collaborative that's demonstrated that is the Keystone collaborative. And also there was a lot of work across the country doing the IHI campaigns where hospitals initially, their compliance with the bundle was low and as they began focusing on it that increased.

But is there hard evidence? Are there hard studies summarizing that? I'm not aware that that data exists.

CHAIR BROTMAN: I think part of the problem is the inconsistency in administering the bundles. Let me go to Aaron first and then I'll go to Peter.

MEMBER MILSTONE: I feel like a lot of the comments that are coming up are really related to reliability and validity so we can probably discuss those in a few minutes. But I wanted to clarify something with either Ed or the developer about the Joint Commission requirements. Because currently the Joint Commission, one of the National Patient Safety

Goals is to reduce central line-associated bloodstream infections. And they require documentation of compliance with best practice.

And I know a lot of institutions have interpreted that as putting a checklist into the medical record. So I don't know if the developer has a sense of how many people -- I know this was someone else by Peter, but how many people are using this as a way to comply with Joint Commission. Because all hospitals are required to document, not just to do this but to document compliance with best practice. And I think a lot of them are satisfying that requirement by using a checklist and either putting it into the paper chart or putting it into the EHR.

MS. JACOBSEN: I would absolutely agree with that, that it is well-utilized and that it has become a very effective tool for Joint Commission review and overall process. Ed, I'd ask you to comment also, please.

MEMBER MILSTONE: But that also gets at the question of is there a performance gap.

CHAIR SEPTIMUS: Yes, this is Ed. Yes, I agree with that in general. I think the question that I have for this particular element is there seems to be in most facilities at least a high level of compliance now with this bundle. And so I think the question that we're at in terms of performance gap, is there still a performance gap. If we had done this 5 to 10 years ago we would be looking at this extremely differently than we're looking at it in 2012. So the question I think for the committee is is there still a performance gap that would require documentation of this bundle.

CHAIR BROTMAN: Peter?

MEMBER HAVENS: So maybe the other way to look at that is to ask yourself the broader population-based question of how many hospitals are doing it and how many aren't.

So that the performance gap is not within a hospital but using the hospital as the unit of measure or the state of Minnesota would look at all the hospitals in the state and what percentage are or are not using the bundle. So then that becomes a -- would be the measure that we would look for here. It's not here but we would identify based on the conversation that there are still hospitals not doing it and leave room for -- that would identify a gap in care and leave room for us to say that yes, there is a performance gap because not everybody is doing it.

CHAIR BROTMAN: Reva?

MEMBER MURTHY: So just to help answer at least one question from one state. I have data from California where CLIP measures have been reported for 3 years and this is data from 2011 of some 400-plus hospitals that are reporting in. These are again self-reported data. And it shows in adult-only ICUs and pediatric ICUs 96 percent

and 95 percent respectively. So in terms of addressing -- with all the limitations of self-reporting that's the -- is that really measuring a performance gap or is it just measuring self-reporting? But that's what's presumably out there. There's no auditing.

MEMBER HAVENS: Okay so potentially the question would be of the total California hospitals what percentage actually reported. And the gap would be in the people who didn't report. And that would be the real target of the measure then.

MEMBER MURTHY: There are actually, of the hospitals there are only four hospitals that didn't report.

CHAIR BROTMAN: Okay, Tiffany.

MEMBER OSBORN: Perhaps, Diane, we can get some more information from the IHI data. So you, I know that you were -- IHI was asked to assist in implementing the bundle. So what was the rate of bundle compliance prior to your work with the hospital system and how many

hospitals did you work with?

MS. JACOBSEN: The, again, the rate of compliance early on when the bundle was developed was very low and over time that increased. So there's several collaboratives, critical care collaboratives, ICU collaboratives that have been in place and then over the period of the 500,000 -- million lives campaign, excuse me, 100,000 and 5 Million Lives Campaign, the increase.

And in the state where reporting is "required" quote unquote or how the public's reporting obviously those rates are -- reporting has increased dramatically. So, it's variable depending upon the way in which you look. All of the data reported to IHI is clearly voluntary.

MEMBER BLANK: I was just going to comment. Early on it was abysmal, the statistics with this measure. Back in experience from Pittsburgh Regional Healthcare Initiative in 2001 when we implemented in 30

hospitals, very low. So it's become a standard of care at least in our neck of the woods and we also have it in our pay-for-performance program monitoring it. So it's close to 100 percent.

Very much like the surgical safety checklist from World Health Organization when we had hospitals start to implement that. Very low. Almost nearly 100 percent right now. So a lot of value in it.

The other comment that I wanted to make and try to get some opinion from Diane on on this is that I do think with the CDC National Healthcare Safety Network that if an outcome -- a CLAB is identified I think they do ask you to identify whether or not they were in compliance with the bundle whenever they inserted it if it was an ICU event.

MEMBER FAKIH: But you know, this can be all done through the procedure note you know. I mean it can be done through a different way.

The other thing with the bundle is the component of daily evaluation which is not part of the checklist. And this is an all or none bundle. You know, it's like yes or no, you have all the elements.

CHAIR BROTMAN: Helen, did you want to make a comment?

MS. BURSTIN: I'd just point out that again I think they're very similar but there's the IHI bundle. Then there's the NHSN bundle which we actually did look at it. They're very, very similar. There actually is published data, I was just pulling it up, on the NHSN compliance with the bundle as of 2010. At least in the 250 hospitals they randomly looked, a cross-sectional study of NHSN hospitals. They found 38 percent reported high compliance with the bundle. Again, that's all comers across NHSN. We can compare those but it also might be helpful to have Diane speak to how the IHI bundle may be different from the NHSN bundle as well.

CHAIR SEPTIMUS: But NHSN is not going to require CLIPs to be reported anymore.

CHAIR BROTMAN: Okay. Let's go to -- Tom, did you have a question? I'm sorry. Mike?

MEMBER FARBER: Just a comment. Again, I think that in this discussion about the bundle I think that the elements of this bundle are what hospitals are expected to do and do measure. In deference to yesterday when we talked about sepsis and the bundle there was considerable concern that some of the components of the bundle really don't need to be there or wouldn't need to be done to have a good compliance. So I think that in this regards as we've heard that the components of the bundle for central line are well regarded, usually detected by chart review and then reporting by the hospital epidemiology.

CHAIR BROTMAN: Okay, I don't want to get too far behind, but let's -- Adam, Tom and Peter and then we've got to go for a vote.

So, quickly.

MEMBER THOMPSON: Yes, I just had two. One comment and a question. The question is concerning whether there are any health disparities around this, whether this is happening in all hospitals or we might be seeing it in areas where their underserved populations might not be using this.

The second comment also has to do with measure, with importance to measure which is when we're working with collaboratives we always tell people even if you reach a high performance if the measure is important and you know that it has a significant outcome that you would continue to measure it. And I think what I'm hearing from around the table is that this has had significant improvements in the reduction of infections. And so it might be that even though performance is high we would still want to measure it because we know that the outcome from it is so good.

CHAIR BROTMAN: Yes, go ahead.

MEMBER FAKIH: You know, the outcome is already mentioned which is central line. So the final outcome is there. The CLABSI, central line associated bloodstream infection is measured. It's something that should be measured for every single ICU. It's mandatory. It's sent to NHSN.

So this is a process measure and it's based on documentation and paper documentation. So that's, I'm not -- again I'm not debating any of the evidence in fact other than a couple that are very tough to get. But you know, as a bundle, just documentation, it doesn't mean it's going to translate into real practice. So what's written may not be what's happening. That's the only thing I'm saying.

CHAIR BROTMAN: Okay. Let me just move this on a little bit. Tom quick and then Peter quicker.

MEMBER GIORDANO: So as a person who does not follow this literature I've got

to get us back to the question here and I'd appreciate an answer from the infection control experts on the panel. Is there a performance gap? Can you have a hospital now in 2012 -- wait, yes -- that is -- that does not measure this because Joint Commission requires it or someone else requires it? Is there a performance gap possible or are you -- if you have a functioning ICU are you going to be at 95 percent or better on this measure? Is there a quick, simple answer to that question?

CHAIR BROTMAN: If anyone has an answer to that question quickly.

MEMBER MILSTONE: Yes -- no, I think one way to think of it is most people are probably using it. Whether there is complete adherence is -- there's no data on. So I think there truly is no data. I think my opinion is that most people probably apply a bundle of some sort because of the national attention, but there's little data on adherence aside to what's being presented through some

collaboratives.

MEMBER GIORDANO: So you could have a hospital that's not doing this. It's possible.

MEMBER MILSTONE: Or that's not doing it well.

MEMBER GIORDANO: Yes. Okay, thank you.

MEMBER OSBORN: That's not compliant.

CHAIR BROTMAN: I think you could have any of the permutations. Peter, you don't need to make a statement? Let's go to the vote now for performance gap.

MS. KAHN: Voting on 1(b) performance gap. You can go ahead and start. We have two high, five moderate, five low and seven insufficient evidence.

CHAIR BROTMAN: Okay, so that's a stop and that fails so we're going to move on.

CHAIR SEPTIMUS: Diane, thank you

very much.

MS. JACOBSEN: Thank you very much.

CHAIR SEPTIMUS: I think in many ways this is a credit to efforts like IHI and others in driving compliance to now that opportunity. So in many ways I consider this a success even though the measure failed.

MS. JACOBSEN: Thanks again. Have a great day.

CHAIR SEPTIMUS: Thank you.

MS. JACOBSEN: Bye bye.

CHAIR SEPTIMUS: Okay, we're going to keep on going, 0405 "Pneumocystis Prophylaxis." Dr. Peter.

MEMBER HAVENS: There is no developer who does this before I do?

CHAIR SEPTIMUS: They're coming back. I'm sorry. I thought we'd beat them up so much before that they had left.

(Laughter)

CHAIR SEPTIMUS: Bob has become Jenna. Forgive me, I'm sorry. I skipped a

step. Go ahead.

MS. WILLIAMS-BADER: Did you just want me to introduce the PCP prophylaxis measure right now?

CHAIR SEPTIMUS: Just any comments you want to make from your perspective and then we'll go through the measure in detail under Peter's guidance.

MS. WILLIAMS-BADER: Okay, great. Thanks. This measure is included in the PQRS program and as we very recently learned I think at the end of last week it has also been included in the measures for stage 2 of meaningful use. Since it is a measure that is included in meaningful use we have an e-measure specification for the measure and that was included in your packet. So that's a little different than most of the other HIV measures.

I do recognize that this is a complex measure because we do have three different denominators to account for the varying indications of PCP prophylaxis for

different age populations. But I would like to point out that when we did the testing of the e-measure among three different sites they all found that the measure is feasible as specified despite the complexity of the measure because the measure does rely on discrete, fairly easy to capture data elements. So I just wanted to make that point. I think that's it. Thank you.

CHAIR SEPTIMUS: Okay, Peter, if you will start off with impact.

MEMBER HAVENS: The impact concerns the concept that HIV is prevalent, that late diagnosis is still common, that CD4 cell counts below 200 continue to occur in the adult population so there is a substantial proportion of people in this country who would still fall into this category even in the era of highly active antiretroviral therapy availability for many people.

The summary statements did not include specific percentages of those sort of

focus points but clearly the data are available in the references that were given.

The complexity of the measurement comes from the different cutoffs for PCP prophylaxis in different age groups. CD4 below 200 is appropriate in use for children age 5 and older. Between ages 1 and 5 the appropriate risk identifier is CD4 percentage of 15 percent. And below age 1 PCP risk is difficult to link to CD4 number or percent. So, prophylaxis is recommended for all children under age 1.

And finally, PCP prophylaxis when used in these risk groups saves lives based on data from randomized controlled trials in both adults and children. So the impact is high and the data are of excellent quality.

CHAIR SEPTIMUS: I think this is pretty straightforward. Unless anybody wants to comment why don't we just vote on the impact. Anybody else? Okay, let's vote. Don't vote yet.

MS. KAHN: Voting on high impact.

Go ahead and start.

CHAIR SEPTIMUS: Go.

MS. KAHN: So you have 19 votes for high, zero for moderate, low and insufficient evidence.

CHAIR SEPTIMUS: Okay. Peter, I think we can go onto the scientific evidence which I think is pretty much consistent with pretty much what you said before. But is there any other comments that you want to make about the science?

MEMBER HAVENS: No. The identified populations PCP prophylaxis saves lives based on data from randomized controlled trials in adults and observational trials in children in the United States and randomized controlled trials in children in other countries.

CHAIR SEPTIMUS: Seeing no hands we'll vote on the evidence.

MS. KAHN: Voting on 18 evidence. You can go ahead and start. Can everyone

press it one more time? We have 19 for yes, the body of evidence meets the guidance, zero for no and evidence does not meet the guidance and no for insufficient information.

CHAIR SEPTIMUS: Okay, the next one is going to be opportunity and gap performance.

MEMBER HAVENS: Section 1b.2 on page 4 of the PDF identifies in 2009 61 percent and in 2010 76 percent compliance with this measure, identifying a gap in care.

CHAIR SEPTIMUS: Any other want to comment on gaps? Okay, well then we vote on -- oh, I'm sorry. Kathleen.

MEMBER BRADY: No, I just wanted to know if there was a breakout by the different age groups for the gap data.

MEMBER HAVENS: It was not supplied here.

CHAIR SEPTIMUS: I guess, again the same question we should weigh about disparities.

MS. WINKLER: Does anybody have

anything to offer from your own personal experience or knowledge on either of those? Okay.

CHAIR SEPTIMUS: Tom?

MEMBER GIORDANO: I don't have any data on disparities for this particular outcome. The gap that you cited is bigger than I would think certainly than what we find in our internal data. What was that, what was the source of that data?

CHAIR SEPTIMUS: I think it's PQRS.

MEMBER GIORDANO: PQRS? I'm surprised.

MEMBER HAVENS: Well, no. So I think this is an important concept. A guy who runs a big well-run clinic is shocked by the size of the reported gap. And this is one of the cheapest, most effective things you can do for people with low CD4 cell count. There continues to be a gap in care. It's important that this measure be adopted broadly if we find it to be a valid and reliable measure of what

we're trying to look at.

CHAIR SEPTIMUS: Adam?

MEMBER THOMPSON: Yes. And one thing about the disparity data and the representatives from HRSA who were at this presentation might also be able to speak to this. But we saw a presentation that did indicate there were disparities in the individuals who were prescribed PCP prophylaxis that was broken down by race and ethnicity with persons of color being less likely to be prescribed PCP. And then it was cited how important this is to get it and the fact that there is disparity in that is I think something that needs to be looked at.

CHAIR SEPTIMUS: Okay. So seeing no other comments let us vote on the performance gap.

MS. KAHN: Voting on 1(b) performance gap. You can go ahead and start. You have 14 high, 4 moderate, zero low and 1 insufficient evidence.

CHAIR SEPTIMUS: Okay. So now we're going to go to reliability and then validity.

MEMBER HAVENS: Concerning reliability on page 9 and 10 of the initial PDF you should note the modification of the measure to allow for no prophylaxis if the CD4 cell count was low on a single measure followed by adequate on the next measure. This has resulted in a change in the measure so that the CD4 is obtained in the first 9 months of the measurement year so it can allow for a transient low followed by a normal.

If we look at the reliability of using automated reporting compared to the visual record inspection reliability seems to be high. In fact there was no difference found in the two measures documented at 2a2.3 on page 12.

This was from a study of 242 patient encounters but I'm not sure how many patients were actually identified in that study done

in the Midwest region in 2009. So it would seem to be reliable although the changes might modify reliability going forward and we would urge users to continue to try to monitor reliability with the changes made.

CHAIR SEPTIMUS: This is again, e-specifications. We're not talking about PQRI.

MS. WILLIAMS-BADER: Well, there is a PQRS measure, right, and then this. We actually have the e-measure here. This is slightly different from the other HIV measures in that. For those we'll need to show you -- bring to you the e-specifications. For this one we do actually have the e-measure already available.

MS. WINKLER: But I think we've already talked about the fact that we only have testing data for the EHR measures so that's really what we're talking about for the endorsement.

CHAIR SEPTIMUS: So what's on page

12 is from?

MEMBER HAVENS: From the EHR review comparing electronic to manual observation there was zero difference in classification, suggesting that you can reproducibly identify what's happened.

CHAIR SEPTIMUS: This is E. This is not PQRI? No.

MS. WILLIAMS-BADER: Correct. That's data from an EHR.

MEMBER HAVENS: But we should point out a weakness as stated on 2b.6 page 14, the reproducibility of the measure has not been measured across data sources. If this is going to be used broadly we would urge users to try to identify reproducibility across data sources.

CHAIR SEPTIMUS: Any other comments?

MEMBER OSBORN: I just wanted to point out from what I'm seeing here and, you know, tell me because maybe I'm missing

something else, it looks like the sample that was 242 patient encounters in one institution. Is that correct what I'm seeing here? Or is there some other data that I've missed?

MEMBER HAVENS: As I read it that was the -- those were the total data presented for reliability, yes.

MEMBER OSBORN: So it's 242 patient encounters.

MS. BURSTIN: It's a network -- PCPI could jump in here -- a network of community health centers in the Midwest with 242 patient encounters.

MEMBER OSBORN: And can you help us? Because I know that I was sort of was discussing this last night before we left, but explain again regarding when we're looking at reliability and validity of testing the measure. Can you just explain to us again what we're evaluating here from that perspective?

MS. BURSTIN: Again this is a bit confusing. I apologize, I think I led you

astray yesterday before Heidi set you straight. So, essentially because it is an automated measure there is an element of reliability that's assumed. So instead what they're really looking at when they do the visual inspection versus the automated results is really the validity of the measure and reliability is assumed in some ways. These are really kind of co-linked for e-measures.

MEMBER OSBORN: So, but we're still, we still comment on reliability as well as validity, is that correct?

MS. BURSTIN: Right, but they really are intermingled concepts. I know Karen Pace is listening in, our methodologist. Karen, anything you want to add?

MS. PACE: This is Karen Pace.

CHAIR SEPTIMUS: Breaking up.

MS. PACE: Is that better?

MS. BURSTIN: Yes.

CHAIR SEPTIMUS: Much better, thank you.

MS. PACE: Okay. As Helen was saying the -- when you get at the data element level of reliability and validity with an automated program you know you're going to get the same results every time which would be the reliability. And so we, the measure testing task force really directed that efforts be placed on data element validity which gets at is the e-measure accurately pulling the correct data. And so when you're at the data element level the reliability and validity are so closely linked and to mitigate some of the burden of testing the measure testing task force really said if you're going to do data element testing to focus on the data element validity. Which in many cases would be comparing the output for example, the numerator, the denominator that's the output from the e-measure specifications to a visual inspection of the entire record to see if the e-measure is really accurately reflecting the data that is in the medical record.

CHAIR SEPTIMUS: Kathleen, did you have a comment?

MEMBER BRADY: Yes and it's related to that. The reliability and validity testing was done at the measure level, not at the individual data element level, correct?

CHAIR SEPTIMUS: Does someone want to comment on that?

MS. CHRISTENSEN: I will if that's okay. Hi, Keri Christensen from the PCPI. We participated in the testing. The analysis that we have provided you is at the measure level. We do look at the data element level if there's concerns at the measure level which there were not for these measures. But we do collect the data for both the data element level and the measure level. And the number of patients is actually double the number that we would need for statistical significance for that particular testing.

MEMBER BRADY: But based on the guidance that we have from you that it can't

-- because it's only at the measure level that what we've been reported it can't be rated above moderate. Is that correct?

MS. BURSTIN: That's correct, although again this is a little bit complicated because this measure essentially is looking at one element, did you get prophylaxis. So they're probably pretty correlated would be my guess. But yes, I think that's a fair assumption, Kathy.

CHAIR SEPTIMUS: Okay, so I think we're ready to vote on reliability I think. So let's get prepared to vote.

MS. KAHN: Voting on 2(a) reliability. You can go ahead and start. So you have 1 high, 16 moderate, zero low and 2 insufficient.

CHAIR SEPTIMUS: Okay. Then let's go to validity. I think we, unless someone has -- we sort of talked about both together so unless there's no additional comments. I don't see any. Let's go --

MEMBER HAVENS: Excuse me, there are comments.

CHAIR SEPTIMUS: Oh, Tom just put his thing up. Thank you.

MEMBER HAVENS: Well, just to review what was here, is the measure valid. The face validity is terrible as cited in the document on page 13 in an incredibly small study which suggests 50 percent face validity. So I think that the data presented here is extremely poor. That's on 3a.2 page 6 -- or no. Well, but that's on page 13, the face validity study in a very small group of people which was evenly split over whether or not the measure is valid. So in choosing studies I think it might be prudent for the developers to choose larger studies that would better support the use of this measure.

I could point out, however, that intrinsic in its wide use and you can see for example on page 16 3a.2 that it's being used by the HIV quality people suggesting that other

groups might consider the face validity much higher than what was reported as supportive evidence for this part.

The measurement validity as we've discussed about was what was reported for what I would consider to be reliability.

CHAIR SEPTIMUS: Okay. I think there was only six in that. It was very, very small but you are correct on what's in the document. Jenna, do you want to respond to that?

MS. WILLIAMS-BADER: Yes, if I could that would be great. Thank you. Yes, it wasn't a study, it was our expert panel were asked to review the face validity for all the measures.

And really the major concern here I think was about the youngest age population and whether or not it's appropriate to just look for the one-time prescription of PCP prophylaxis among the much younger age group because the evidence I believe says that they

should be on it for a longer amount of time, for a certain period of time.

So as far as the older populations there really was not any concern among our experts about the face validity for those older populations. I think it was really just about that younger population where they had the concern.

And as you can see when the three test sites were asked about the face validity for this measure they rated it very highly and like I said, it has been chosen as one of the measures in meaningful use which I think indicates that others believe this is an important measure.

CHAIR SEPTIMUS: Aaron?

MEMBER MILSTONE: I have another EMR question for you. So, I was looking at the logic for -- because the way you list the denominator is by -- it has a group category as medication for PCP prophylaxis. So I was trying to figure out how that's going to be

captured using different EMRs. And I'm looking down now at pages, I don't know if it says down in your logic where it lists all the different codes. The categories include value set name and then there's one that says code and then there's one that says descriptor. And in the field for value set name there are a bunch that are listed as pneumocystis, PCP prophylaxis and then under code there's a number. I'm not sure if that's the CPT code that you refer to. And then the next one is the descriptor. It lists things like batch and other drugs.

So I guess my question again is using an EMR that has these data fields this should be reliable, right? You're going to run it at the same time, get the same thing every time and the validity should be good because you have all the drugs listed here that should get pulled.

But I'm just wondering whether we know that there are other EMRs that have similar

codes, whether these are, as you mentioned before, free text fields or people are shaking their heads so jump in. That's why -- I'm asking how this would compare to different EMR in terms of the ability to quickly identify the drugs.

MS. WILLIAMS-BADER: I can definitely answer that. During our feasibility testing for these measures which was really to see whether there are standardized structured fields for these data elements, the test sites found that these are all available in structured fields. So it should be similar across all EHRs that at least the test sites where we tested them did have that in structured data fields, not in free text notes or other types of non-structured data fields.

The codes we provide are the codes for -- so we provide RxNorm codes for the medications which is in compliance with CMS's blueprint about which vocabulary you should

use for this particular type of data element. If the EMR is not using RxNorm codes itself they can map to the RxNorm codes. And many EMRs are actually using local codes for certain data elements but that doesn't mean that they can't and shouldn't be mapping to the codes that are provided with the e-measure.

CHAIR SEPTIMUS: One thing I found out about Aaron is that he's a geek.

(Laughter)

MEMBER MILSTONE: I work with a lot of electronic data and with different systems so I understand the difficulties of trying to merge them. So I just want to make sure, as Tiffany said before, that for clinicians that are doing the right thing I want to make sure they're not going to get dinged because it's not getting picked up. So thank you for that clarification.

CHAIR SEPTIMUS: I'm teasing, Aaron. So since this has been deemed meaningful use there will be an incentive to

people to map and to use the standard vocabulary.

MS. WILLIAMS-BADER: Absolutely.

CHAIR SEPTIMUS: Which I must say is a challenge out there. Any other comments about validity? Then let's vote.

MS. KAHN: Voting on 2(b) validity. You can go ahead and start. You have 2 high, 15 moderate, zero low and 2 insufficient evidence.

CHAIR SEPTIMUS: Okay. Feasibility and usability. Peter, let's start with feasibility.

MEMBER HAVENS: It has been in use for a number of years. NQF has asked for input on problems with usability and has acted on issues addressed by different groups which should only increase its usability in the future.

CHAIR SEPTIMUS: Comments from? Okay, we'll vote on usability.

MS. KAHN: Voting on usability.

You can go ahead and start. One more time. You have 10 high, 9 moderate, zero low and zero insufficient information.

CHAIR SEPTIMUS: Okay. The next element is feasibility. Goes into electronic sources, inaccuracies or intended consequences. Peter?

MEMBER HAVENS: The feasibility is high in places where hospital programmers will program this into their medical record so that it can be used. Feasibility is low if you can't get programming back up to do this. The fact that it's been put into meaningful use will be potentially useful if it will open up IT resources at local sites to get it programmed. So the feasibility is potentially well without money put towards the process. But since money has been put potentially better.

CHAIR SEPTIMUS: Tom?

MEMBER GIORDANO: Can I just clarify that what we're talking about in this measure in both feasibility and usability is

the electronic version, not the CPT category 2 code that actually is listed in the document, right?

MS. WINKLER: Just as we did yesterday with the hep C measure that's going to be amended.

MEMBER GIORDANO: Okay. So feasibility there in that situation, I agree with Peter, seems reasonable.

MEMBER HAVENS: It wouldn't be reasonable -- the CPT-2. Help me understand what the difference there is since I'm not that kind of coding --

MEMBER GIORDANO: Well I'm certainly not a coding monster.

(Laughter)

MS. BURSTIN: So, essentially a CPT-2 code allows a clinician to self-attest to the results of what happened during that encounter to answer the measurement question. Since many of these measures haven't even been in PQRS and they don't have data from PQRS

there's no way for them yet to actually assess the reliability of that coding which is self-attestation. So some have actually argued that do you actually need to test what was an attestation. But again, for now they're not on the table.

MEMBER GIORDANO: So my answer to that would be I find that completely not feasible for routine care, that clinicians are going to go in and start coding all these things they said they already wrote down in their note.

MS. WINKLER: Just to keep it real clear all we're looking at here is the EHR specifications for the measure.

CHAIR SEPTIMUS: Any additional comments? Seeing none we'll vote on feasibility.

MS. KAHN: Voting on feasibility. You can go ahead and start. You have 3 high, 15 moderate, zero low and 1 insufficient information.

CHAIR SEPTIMUS: Then the last element of course is the overall suitability for endorsement.

MS. KAHN: So does the measure meet NQF criteria for endorsement, yes or no. You can go ahead and start. You have 18 yes and 1 no.

CHAIR SEPTIMUS: So the measure passes. Jeff, I'm going to give you an alert. We're taking a 10-minute bio break.

MEMBER BEAL: Thank you.

(Whereupon, the above-entitled matter went off the record at 10:03 a.m. and resumed at 10:16 a.m.)

CHAIR SEPTIMUS: Okay, let's settle in, folks. Operator, can you tell us who's on the line, please?

OPERATOR: I'm showing that we have Karen, John and Jeff online.

CHAIR SEPTIMUS: So no one from the CDC has called in?

MR. BROOKS: This is John Brooks.

I'm here.

CHAIR SEPTIMUS: Okay, John, thank you. Okay, we're getting ready to start.

MR. BROOKS: Sure. I'm just going to listen in mute mode until -- I'll try again if I need to say anything or if somebody asks a specific question.

CHAIR SEPTIMUS: Okay, thank you.

MR. BROOKS: You bet. Thanks.

CHAIR SEPTIMUS: Because the next one is a HRSA measure, 2083 "Prescription of HIV Antiretroviral Therapy." So we'll let our developers make a brief intro.

MS. MATOSKY: Good morning, everyone. My name is Marlene Matosky. I'm from HRSA's HIV/AIDS Bureau.

And I'd like to just say that I am joined by an esteemed group of colleagues who are part of our measurement development team. The table apparently is not big enough for all of us and we were the two that had to come up here by ourselves but we have folks from

our team, from CDC and HRSA. We have somebody who works out of the Secretary's office out of HHS with us. And we have folks from CDC on the phone also.

So I just wanted to say that this project was a significant experience for us as measure developers. We are here in a very different way in that we're not here for maintenance of measures, we are here for initial endorsement. So I hope that you could take that into consideration as we're moving forward.

We feel that folks here at HRSA, CDC and HHS are very well positioned to be stewards for measures because in many respects we are seen as the experts and the go-to folks within the field of HIV. We fund within HHS a significant portion if not all of the publicly funded services related to HIV care, treatment and prevention. And saying that we know that we will have a significant impact in not only the usability and the feasibility and the

in-field implementation of these measures.

We see these measures not only being used within the HRSA programs, we also see these measures being used at the HHS level and public reporting programs also. Three of the five measures we're bringing to you have been endorsed by the Secretary of Health and Human Services, so Dr. Sebelius is behind and has endorsed these measures. So they would have broad applicability across federal programs.

Thinking in general about performance measurement we see performance measurement as just one side of the coin. We see the other side of the coin as quality improvement. We're not in the business of measuring things just to measure things. We hopefully -- and our intent is that we will see quality improvement.

As many of you know there are significant disparities unfortunately within HIV care, treatment and prevention and these measures are well designed to point these out.

I'm just checking my notes here. And I think that's all I have. Is there anything else you would like to add, Dr. Cheever? Thank you.

CHAIR SEPTIMUS: So, we're dying to hear from you.

MEMBER ELAM: Thank you. So as was just stated this is measure 2083. It is a new submission. It's a process measure. It's titled "Prescription of HIV Antiretroviral Therapy."

Brief description of the measure. It's the percentage of patients regardless of age with a diagnosis of HIV prescribed antiretrovirals for the treatment of HIV infection during the measurement year.

The numerator is the number of patients from the denominator prescribed HIV antiretroviral therapy during the measurement year. The denominator is the number of patients regardless of age with a diagnosis of HIV with at least one medical visit in the measurement year. There are no patient

exclusions.

The data source is electronic medical records, electronic clinical data, pharmacy and paper medical records. The level of analysis is clinician group, practice, community, country, city population, regional and state.

So looking first at impact our work group consensus was that this was high-impact. Ongoing evidence about HIV shows that it's a communicable infection that leads to a progressive disease with a long asymptomatic period. Fifty thousand plus or minus new infections per year in the United States. Without treatment most persons develop AIDS within 10 years of infection. Antiretroviral therapy delays this progression and increases length of survival.

ART reduces HIV-associated morbidity and mortality by maximally inhibiting the viral replication. Durable viral suppression improves immune function and

quality of life. It lowers the risk of AIDS defining and non-AIDS defining complications and prolongs life.

There's emerging evidence that also suggests additional benefits of ART-induced viral load suppression include reduction in HIV-associated inflammation and possibly its associated complications. And measures of viral replication can predict HIV disease progression among untreated HIV-infected time to clinical progression and mortality is faster in those with greater viral loads.

And then last, antiretroviral therapy has also been shown to reduce transmission of HIV. The risk of sexual transmission is highly correlated with HIV viral load in blood and genital secretions of the infected person and antiretroviral therapy reduces viral load in blood as well as viral shedding in body fluids including the semen, cervico-vaginal and anal-rectal secretions. So basically improved treatment equals

decreased viral load equals decreased transmission, morbidity and mortality.

One of our work group members did make mention as far as the impact on this that the -- there was insufficient information that the measure did not show deficiencies in ART prescriptions. So any questions about impact?

CHAIR BROTMAN: Okay. I think that -- thank you for that great summary. I think this is fairly straightforward. If there's no discussion let's go for voting on impact.

MS. KAHN: Voting on high impact. You can go ahead and start. Can we have everyone press it one more time? So 18 high, 1 moderate, zero low and zero insufficient evidence.

CHAIR BROTMAN: Okay, that overwhelmingly passes. Let's go onto the evidence.

MEMBER ELAM: So with regards to quantity of evidence there were greater than

five studies cited. These included randomized controlled -- or randomized clinical trials, meta analysis and observational studies. Several of those observational studies were a collaboration of cohort studies.

The type of evidence was based on clinical practice guidelines. The HHS guidelines cited recommendations for use of antiretroviral therapy in HIV-infected adults and adolescents to reduce associated morbidity and mortality and reduce the transmission of HIV.

The HHS guidelines in pediatric HIV-infected populations highlight that ARVs are associated with enhanced survival, reduction in opportunistic infections and other complications, improved growth in neurocognitive function and improved quality of life in children.

A work group concern was that this measure basically incorporates all ages for treatment. And the comment was while we

recognize the importance of this clinically the current guidelines that are presented for the pediatric population in children less than 5 years of age state for those that are asymptomatic with a CD4 percentage rate of 25 percent and a viral load of less than 100,000 copies, a physician should consider treatment.

Quality of evidence, body of evidence used for the recommendations on treatment to reduce HIV-associated disease and death as a whole. The quality of the RCTs was high. Intervention and control groups had similar baseline characteristics and retention rates were high.

Observational studies were large and used advanced statistical methods to minimize the bias and confounders that arise when observational data are used to answer questions about when to initiate treatment. Nonetheless there were unmeasured confounders which may have -- affect these analysis. And

the consistency of the evidence, effect on disease progression by pre-treatment CD4 count, very consistent findings and narrow confidence intervals in the majority of studies for those with CD4 counts of less than 350.

The CD4 count of 350 to 500 shows statistically significant impact on disease progression, death and consistent magnitude of impact hazard ratio of 1.3 to 1.7 and narrow confidence intervals.

The CD4 above 500, data is less strong. There's no impact on progression to AIDS or death. And a work group comment was the intent of -- for treating over 500 CD4 count is that one may treat. And it was noted that in large jurisdictions including San Francisco and New York City health officials are implementing policy that all patients diagnosed with HIV regardless of CD4 counts are being treated. Work group members are uncomfortable being held to a standard backed by limited evidence.

On the whole the results were generally consistent within categories and the impact of treatment decreased as pre-treatment CD4 count increased.

There was also information about effect on transmission. Large random controlled trials of serodiscordant heterosexual couples documented a 96 percent reduction in risk of transmission for the treatment group compared with the deferred treatment group. And studies show an association between plasma viral load and heterosexual transmission.

Work group comment on this was that there's insufficient data to require treatment of all patients with HIV. This does not provide exclusions for patients that refuse treatment or are not prescribed treatment for various reasons.

CHAIR BROTMAN: Okay. Any discussion on the evidence points? Aaron, did you want to talk about the pediatric issues?

MEMBER MILSTONE: Yes. Maybe I should defer this to Peter since he treats more peds HIV I think than I do, but I guess I just have trouble because again there's no evidence in children over the age of 5 who have higher CD4 counts. So I think this is a great measure, I think it's very important but there's no evidence and it's not the current state or recommended. I think we're moving in that direction but it's not the current standard. So I have trouble with the measure as encompassing all patients with HIV as opposed to maybe a population of greater than 13 years of age where it's more the standard.

And I think when I think of pediatrics, you know, most of the children we're seeing now are in the adolescent world. There are a lot of adolescents that have trouble with adherence to medications who may have higher CD4 counts who are watched because of concern for compliance. I think that's why there's some question amongst experts.

But I don't know if Peter wants to comment. But I have trouble with the evidence as a whole because there's a population that it doesn't include.

CHAIR BROTMAN: Peter?

MEMBER HAVENS: Thank you and as a disclaimer I'm on that guidelines group.

I like this, the simplicity of this approach to measurement and think one of the questions that is inherent in the current discussion is how will the data be used. So if -- I think it would be useful data to be able to document whether 50 or 80 percent of children are being treated independent of whether the guidelines say do it or consider it. There are important issues related to the potential public health impact of treatment in sexually active adolescents and adults which don't pertain to children. Therefore the balance of immediate treatment in children depends completely upon proven benefit versus potential for toxicity of long-term drugs and

does not have perhaps the extra benefit of the potential for public health impact by bringing down general secretion virus load and reducing transmission. So that is an important reason that the pediatric guidelines are different.

But even though those guidelines say consider instead of do and depending on how you read the adult guidelines you could consider rather than do. I think a measure of current practice that this allows is an important consideration. So I'm very supportive of this approach to it.

CHAIR BROTMAN: Kathleen and then David.

MEMBER BRADY: So I mean, I think there's going to be a lot of discussion regarding the 500 CD4 count and above. And so I mean even within the guidelines themselves for adults it's a B-3 recommendation which is moderate and based on expert opinion.

So, but on the other side of that

there is data, actually I mentioned this data before, an analysis that CDC has done. And I don't know if John Brooks could add to this, but an analysis looking at the Gardiner cascade using surveillance and the data from the Medical Monitoring Project. It was determined that by changing the guidelines from less than 500 to over 500 the overall impact that would have on the number and percent of people receiving antiretroviral therapy would be 3 percent. So you know, we're talking about you know a small number of people that are going to be included in this where we're questioning whether it should be in there or not.

CHAIR BROTMAN: Thank you for bringing that up. David.

MEMBER SPACH: And I just wanted to clarify I think the subtle shift that has occurred in the last year regarding the guidelines and the above 500. Previous to the most recent guidelines it was recommended to consider therapy with patients with CD4 count

above 500. In the HHS guidelines most recently it recommended for all patients. It's just the strength of the recommendation as Kathleen nicely outlined is a B-3 recommendation.

And also, the other major, widely viewed guidelines, the International Antiviral Society USA guidelines came out this summer. They also recommended treatment for all patients. So, and there is some albeit not randomized controlled trial but the NA-ACCORD study suggested a survival benefit in people above 500. The HIV-CAUSAL study suggested a morbidity benefit in patients above 500. So I think this is a controversial area but the major experts around the country that reviewed this in the most recent guidelines both IAS USA and HHS recommended treatment for all patients regardless of CD4 count.

CHAIR BROTMAN: Thank you. Tom and then we'll --

MEMBER FILE: Okay, thanks. Well, just a couple of things about the lack of

exclusions. Number one, I think this is going to be a big issue for disparities. I mean we have lots of patients who are on the Ryan White waiting list and depending upon their CD4 count and their clinical status may be on the waiting list for over a year before they have antiretroviral therapy.

And then secondly we have lots of patients who like what you were talking about with compliance are just not ready to start, yet we can tell from compliance issues. And you know, and the clinical -- won't have high CD4 counts. We sort of wait and counsel them. And so I was just going to say there are some exclusions here that I think are valid.

CHAIR BROTMAN: And that goes to the performance gap that we're going to get to as well. Doug.

MEMBER CAMPOS-OUTCALT: So what I'm hearing people say is that while these groups are very authoritative and expert the evidence is -- or the basis for these

recommendations currently is mostly expert opinion. Or not?

MEMBER BRADY: Only for the persons who have a CD4 count above 500. The data is very clear for persons who have a CD4 count below 500. And what I was saying before, the number of people who actually have a CD4 count above 500 who present at time of diagnosis is extremely small. I mean nationally you know over 30 percent of people who are diagnosed with HIV have an AIDS diagnosis within 12 months. And the data regarding, you know, if you do have a CD4 count above 500 at the time of presentation, the overall time period where you would wait where you would meet that less than 500 designation was less than 12 months. So we're talking about, you know, initiating therapy very soon in most of these of people who were above 500 anyway.

MEMBER SPACH: And transmission benefit.

MEMBER BRADY: And right. And

that's actually one of the things that was not taken into account in terms of those, the new treatment guidelines is that there is a 96 percent reduction in transmission of HIV in people who have a discordant partner.

MEMBER CAMPOS-OUTCALT: Do you think that the recent emphasis on increased screening will affect that? Those numbers, in other words the percentage appearing with 500 above or below.

MEMBER BRADY: It hasn't so far. I shouldn't say that entirely. That's -- in some jurisdictions it has but in general. In D.C. it has made a big difference although I kind of question their data to some degree. But for the most part that's not been shown nationally.

CHAIR BROTMAN: Okay. Mohamad?

MEMBER FAKIH: Just a question about how, you know, we are focusing too much about the inclusion of those that are above 500 and whether we should treat them or not.

I see this measure as just looking at improvement over time. And you know, we don't have to get into the 100 percent compliance but an improvement say from 40 percent on antiretroviral therapy to 60 percent, that would be very, very -- I'll be very happy with that.

CHAIR BROTMAN: Okay. The measure developer has a comment?

DR. CHEEVER: I just wanted to make a couple of quick questions. One, in terms of children less than 5 I think that's like 0.1 percent of the population in the United States which is part of the reason when we were developing we didn't consider that as an exclusion because it wasn't a large enough -- less than five infected? Oh, okay. So just it's a small number of kids hopefully if we're doing our job on the front end.

Second, in terms of the ADAP waiting list I think that is a valid concern. We do work closely with states to make sure that the

people on the waiting list are actually on antiretrovirals through pharmacy assistance programs. And our survey of states generally confirms that, that everyone that wants to be on drug is on drug.

And the third point which I think the previous speaker just got to around refusal is that we do expect there to be refusals. We don't expect this to be 100 percent. That would look like coercion actually if it was 100 percent in most clinics. But I think that we do see clinics where there's a 50 percent refusal rate in certain minority populations and other clinics where there's a 10 percent refusal rate. So we as the federal government working with disparities in populations would expect that if you have a 50 percent refusal rate there's an issue in your clinic that you need to address. And so we'd want to be looking at that from an improvement perspective.

CHAIR BROTMAN: Thank you. I want to wind this up so Adam, quickly and we'll get

to --

MEMBER THOMPSON: Yes. I just wanted to also add the perspective of individuals who present with over 500 and are not refusing care. One of the added advantages, and there's a lot of us who presented. I had a CD4 count of 860, chose to start medication because the active engagement with my disease every day was the choice to fight it and not wait to get sick. So there's a mental health aspect to it as well as a retention aspect.

And my concern is if you make, if you don't say over 500 is the possibility providers try to deny us that medication. And I know I'm just one patient but my CD4 count has not dropped beneath 1,500 since that day even as an active drug user at the time.

CHAIR BROTMAN: Okay. Curtis, I think?

MEMBER COLLINS: Yes, just a point of clarification. This is for all patients.

So from, you know, above 500 and including below. If we're holding ourselves to the standard that we've held for other measures on this you know I think it's somewhat clear that for the entire measure as a whole that there may not be this level of evidence. Now, has there been discussion about breaking this out? Perhaps limiting it to 500 or under, altering it in some way. You know, I don't know.

And then another question on the greater-than-500 population. If it is indeed 3 percent has there been any cost-benefit studies done on those patients for this measure? That could potentially affect, you know, a large number of patients here. I'm just wondering about the evidence there.

CHAIR BROTMAN: Okay. Kathleen, Tom and Doug and then I think we're going to have to vote.

MEMBER BRADY: I just want to make a comment about the ADAP waiting list. I know

that that's an issue in some jurisdictions but you know, it was recently announced that there is going to be additional funding to try and clear all ADAP waiting lists you know if that passes. But I feel like we should be treating people based on guidelines and not on whether there's an ADAP waiting list. And so.

CHAIR BROTMAN: Thank you. Tom?

MEMBER GIORDANO: On the evidence for persons with less than 200 it's extremely strong, as strong as anything we've looked at in the last 2 days. If you're looking at the 200 to 350 level it's also I would say very strong, again, maybe as strong as anything we've looked at in the last couple of days. And the only issue is -- and the 350 to 500 it's strong. The only issue is this small portion that's greater than 500. We're being asked to sort of assimilate that into an overall summary of the strength of the body of evidence. There's no formula we can apply to get there but in my head it's at least moderate because

you've got very strong evidence for the very largest population that this measure would affect.

CHAIR BROTMAN: All right. Well thank you for that summary. I think at this point let's vote on evidence.

MS. KAHN: Voting on 18 evidence. You can go ahead and start. Everyone press it one more time.

CHAIR BROTMAN: Okay. So it passes.

MS. KAHN: You have 14 yes, the body of evidence meets the guidance, 3 no, the evidence does not meet the guidance and 1 no, there's insufficient information.

CHAIR BROTMAN: Okay. So that passes. Let's address the performance gap next.

MEMBER ELAM: So looking at the performance gap there's considerable variation in less-than-optimal performance across providers and populations.

The data that was submitted referenced three different studies or data sources, the first being the CDC's Medical Monitoring Project which indicated in 2009 that 89 percent of adults, and that's 18 years or greater, had been prescribed ART. Of these, 77 percent had a suppressed viral load at their most recent test and data from the same system also indicate that among all persons in care only 72 percent achieve viral load suppression.

In an analysis of surveillance data from King County, Washington Dombrowski, et al., found that among persons with at least one viral load reported in 2009 65 percent had undetectable viral load at the time of last report. And among persons with at least one viral load reported in 2009 those engaged in continuous care were more likely to have virologic suppression, and that was 69 versus 58 percent. And those that were engaged in continuous care had a lower mean viral load than those that were not engaged in continuous

care.

And the third data source was Kaiser's HIV Challenge 2011 year-end report. Of all members with known HIV infection and on ARVs 94.5 percent achieved viral suppression in 2009. Of all HIV-positive patients in Kaiser Permanente in 2009 69 percent achieved viral suppression, pointing to the need for further improvement across the spectrum of care.

Disparities by population group data, those were addressed in this measure. Gender, race, age, education and income were all cited in the data.

CHAIR BROTMAN: I'm not sure all of those address the actual performance gap but is there any discussion among the work group members that they want to bring up? Tom? David?

MEMBER SPACH: Just real quickly. Hall presented data at the International AIDS Conference that clearly showed a gap in

African-Americans having lower levels of suppressed HIV RNA levels and lower percentage of African-Americans who were on antiretroviral therapy. So there is a gap that's been shown, a racial gap. An ethnic gap.

MR. BROOKS: If I can just interject. John Brooks. That same analysis also showed a gap by age.

CHAIR BROTMAN: Do you know what the statistics are on that?

MR. BROOKS: I'd have to download the presentation but we can get it.

MS. VIALL: And I don't have them from Irene Hall's presentation but I have them from Jacek Skarbinski's presentation on MMP data from CROI 2012.

What we found is that while 89 percent of people living with HIV in care have been prescribed ART based on MMP data. The percentages range when you look at different populations. So it ranges, for age it ranges

from 72 percent among people 18 to 29 years of age to a high of 92 percent for people over 50.

We also found that non-Hispanic blacks are significantly more likely than whites to have not been prescribed ART. We also found that people with CD4 counts above 500 are significantly less likely to be on ART, 66 percent for people with CD4 counts above 500. Eighty-one percent for people with CD4 counts between 200 and 500, and 95 percent for persons with an AIDS diagnosis. In a multivariate model of factors associated with prescription of ART we found young age, so 18 to 29, non-Hispanic blacks, women who have sex with men and persons more recently diagnosed with HIV were less likely to be prescribed ART. And these come from our 2009 MMP data collection cycle.

CHAIR BROTMAN: Thank you so much. I appreciate you filling in a couple of gaps there.

MEMBER BLANK: I was just going to ask, I'm not hearing any literature describing the gap for the less than 18 year age population.

MS. VIALL: That -- MMP is actually restricted to persons 18 years and over.

CHAIR BROTMAN: Okay. If there's no more discussion -- no. If there's no more discussion let's vote on performance gap at this point.

MS. KAHN: Voting on 1(b) performance gap. You can go ahead and start. You have 7 high, 10 moderate, 1 low and 1 insufficient evidence.

CHAIR BROTMAN: Okay. So that passes. Let's go onto reliability.

MEMBER ELAM: So, with regards to reliability there were precise measure specifications in that the numerator was the number of patients from the denominator prescribed ARVs during the measurement year. The measurement year is a consecutive 12-month

period. Numerator details to be included were patients that were prescribed antiretroviral therapy during the measurement year and antiretroviral therapy was described as any combination of HIV medications other than the regimens or components identified as not recommended at any time by the panel on ARV guidelines for adult and adolescents.

The denominator was number of patients regardless of age with a diagnosis of HIV with at least one medical visit in the measurement year. And denominator details to be included: patients must meet all of the following conditions or events. Number one, patients of any age during the measurement year; two, patients diagnosed with HIV during the first 3 months of the measurement year or prior to the measurement year; and patients who had at least one medical visit during the measurement year.

There was no adjustment for -- there was no risk adjustment and no risk

stratification. And the reliability testing, the type of score was better quality equals higher score and it's based on rate and proportion.

Comments from the work group regarding reliability. What about the exceptions that are not accounted for? And I think the other two go to validity.

CHAIR BROTMAN: Any discussion on the reliability aspect of this? Go ahead, please.

MEMBER GIORDANO: Is this measure through electronic health records then?

MEMBER ELAM: Yes.

MS. MATOSKY: So I'd like to clarify that. We did not, as with the other measures that have been presented thus far, specify this measure for use in electronic health record. Rather, we used the HIV Research Network data set and that original data comes from a variety of sources across the 18 sites. They could have paper charts,

electronic health records and such. And they at the site abstract that data and send it to Hopkins, that's the data coordinating center. But the original data came from a variety of sources, whatever was used in the clinic. And we did a testing on the data that came from the Research Network.

CHAIR BROTMAN: Okay. If there's no other comments let's go to vote on reliability.

MS. KAHN: Voting on 2(a) reliability. You can go ahead and start. You have 2 high, 17 moderate, zero low and zero insufficient.

CHAIR BROTMAN: Okay. So that passes. Let's go into validity.

MEMBER ELAM: So I think validity basically is, the response is that it was -- face validity and because it's electronic health record the reliability is moderate to high. Moderate, actually. Moderate.

CHAIR BROTMAN: Okay. Any other

discussion on this? All right, let's vote on validity.

MS. KAHN: Voting on 2(b) validity. You can go ahead and start. Can we have everyone try it again, please? We have 1 high, 18 moderate, zero low and zero insufficient.

CHAIR BROTMAN: Great, that passes. Let's go onto usability.

MEMBER ELAM: This is a meaningful, understandable and useful measure. The HHS work group saw utility in publicly reporting this data. The only concern was that the process for reporting was not outlined.

CHAIR BROTMAN: Okay. Any discussion? All right, let's go vote on usability at this point.

MS. KAHN: Voting on usability. You can go ahead and start. You have 7 high, 12 moderate, zero low and zero insufficient.

CHAIR BROTMAN: Okay. Great. Let's go onto feasibility.

MEMBER ELAM: So the feasibility

of this, the work group had some concerns about the list of the ARVs and potential for difficulties in data collection. The work group would prefer outlining the medications that should not be used together rather than the approach of an abstracter trying to review the regimens to see if they are consistent with the guidelines.

CHAIR BROTMAN: Any comments on the feasibility aspect? All right, let's go to a vote on feasibility then.

MS. KAHN: Voting on --

CHAIR BROTMAN: Do you want to make a comment?

MEMBER MILSTONE: I just wondered if there was a response from the developers on the question.

MS. MATOSKY: Can you just state the question again?

MEMBER ELAM: The work group thought that there was some potential for difficulties in data collection and they were

recommending that outlining the medications that should not be used together rather than the approach of the abstracter trying to review regimens to see if they're consistent with the guidelines.

MS. MATOSKY: The way we were -- the way we intend the measure to be used is that we are going to define antiretroviral therapy as any regimen combination that is not not recommended which is I think what you're suggesting.

CHAIR BROTMAN: Okay. Let's go to a vote on feasibility then.

MS. KAHN: All right, voting on feasibility. Go ahead and start. We have 2 high, 17 moderate, zero low and zero insufficient.

CHAIR BROTMAN: Okay. And finally, suitability for endorsement. Let's take a vote.

MS. KAHN: Okay, does the measure meet NQF criteria for endorsement, yes or no.

We have 18 yes and 1 no.

CHAIR BROTMAN: Great, so that passes. Thanks. Okay, the next measure is actually very close to -- oh, Aaron.

MEMBER MILSTONE: I just wanted to make one comment for the developer. So I was looking at your data table and it lists the drugs and their trade names, but it would be helpful somewhere I think with the measure just to have that list clear as to what the combinations are, the combinations that wouldn't be accepted. I didn't see those. But I assume that changes over time so that will be a thing that develops as you go, correct?

MS. MATOSKY: Those are consistently listed within the guidelines. It's usually like Table 7 and 8 in both the adult and the pediatric guidelines. They're fairly stable tables in that they don't change very often and these are the absolutely never, you know, write that prescription for these

medications. And many pharmacy programs actually query for these when the person takes their prescription in.

So that's why we went that route rather than going you know to program every potential combination of HIV antiretroviral therapy people could be on because we know that there's the first line, then there's the preferred, and so on and so forth. And if you have 20-some odd medications and it becomes ART after awhile the number of possible combinations can become limitless.

And we're very fortunate, as more medications come down the pipeline this measure would be up for regular maintenance in terms of e-specification. And we felt by going the inverse route it would be more stable over time. Thank you.

CHAIR SEPTIMUS: Okay. Next measure actually has a lot of similarities. It is an NCQA. Kathleen is going to present it but we have our fearsome duo back up at the

table. So whether Bob or Jenna would like to make a brief comment and then Kathleen will introduce it to the committee.

MS. WILLIAMS-BADER: Great, thank you very much. Yes, this is a very similar measure in that we're looking for patients who are prescribed potent ART.

We -- when we reviewed the measure with our experts both when it was originally developed and when we recently reviewed the measure and looked at current guidelines to update the measure our panel did decide to stick more closely to what has received strong recommendations from -- in the treatment guidelines. So you'll see that our denominator here are sticking to items that received an A1 or A2 recommendation in the treatment guidelines. We have patients with CD4 count less than or equal to 500 cells. We have patients who have an AIDS defining illness and patient pregnant -- or pregnant patients. Sorry. That's it.

CHAIR BROTMAN: Okay. Let's start with the presentation if we can go to impact.

MEMBER BRADY: Okay, so this -- the title of the measure is "HIV/AIDS Adolescent and Adults Patients Who are Prescribed Potent Antiretroviral Therapy." So in terms of the differences with the previous measure we're only talking about those over the age of 13 as previously mentioned and also those with a history of a CD4 count less than 500 unless they've had an AIDS defining illness. And all pregnant women regardless of CD4 count or age.

So the denominator is all patients age 13 years and older with a diagnosis of HIV and AIDS with at least two medical visits during the measurement year with at least 90 days between each visit who had a history of CD4 count less than or equal to 500. So there's a visit requirement here that is also not -- at least two visits where in the previous measure it was at least one visit. All

patients age 13 years and older with a diagnosis of HIV/AIDS with at least two medical visits during the measurement year with at least 90 days apart who had an AIDS defining illness. And same think for pregnant women, you had to have two medical visits during the measurement year.

And so I feel like we've talked about the impact of this indicator previously. I don't know if there's anything that I want to add.

CHAIR SEPTIMUS: Okay. Also I've been told that Ray is on the phone. Is that right, Ray?

MEMBER BRADY: Yes and actually I did have one comment, and that was actually from the HIV Medicine Association. They actually recommend deletion of qualifications to measure percentage of all patients prescribed antiretroviral therapy. So it would be actually percentage of patients with HIV/AIDS with at least two visits during the

measurement year with at least 60 days or whatever interval is selected for the medical visit measure between each visit who were prescribed potent antiretroviral therapy.

CHAIR SEPTIMUS: Okay, we'll get back to the developer in just a second. Ray, are you on the phone? He's on the webinar, okay. Just, sorry. Okay, any comment from the developers on the HIVMA recommendation? Do they have it? Do they have what you just read?

MS. WILLIAMS-BADER: I'm reading the comment here right now. Deletion of qualifications, I'm actually not quite clear what that means. Judy Aberg is one of the experts though who is on our expert panel and like I said we did revisit the measure with our expert panel and asked them if they'd want to expand this to all patients and not just patients whose CD4 count was below or equal to 500, or pregnant, or patients with an AIDS defining illness. And our panel very strongly

believed that they wanted to have those qualifications in there, that they didn't just want it to be for all patients.

CHAIR SEPTIMUS: Let me ask you this, Kathleen. Does that in and of itself influence the impact of the measure?

MEMBER BRADY: Yes, I would say so because I think automatically if you are limiting it to people who have two medical visits that you're going to be limiting the population to people who are receiving a higher level of care already and not all patients with HIV. I mean, it goes to some degree about retention. So you're only measuring antiretroviral therapy in people who have good retention.

CHAIR BROTMAN: Tom?

MEMBER GIORDANO: I agree with Kathleen's summary of what that means. I'm not sure I agree with her interpretation though. I think it depends on what you want to measure. If you want to measure among

people who are in care to a certain degree what percent are prescribed ART when it's indicated, this would be the measure to do that. If you want to measure among our entire clinic population what percent are prescribed then the previous measure would be better at that. So I think there's -- but I think both would have very high impact because the data are so strong that people with HIV need ART.

CHAIR SEPTIMUS: And that was actually what I was trying to get at. I mean, in other words would this change our vote on the impact of the measure. That's -- Tiffany, did you want to say something?

MEMBER OSBORN: It was really in reference to what you said. I don't take care of clinic HIV patients so you could help me understand this a little bit. But if this is quality measures by which we are holding physicians accountable do we want to hold them accountable for people who don't come back? That's what I'm -- I mean, or unless there's

a system that we are trying to bring them back.

MEMBER BRADY: Yes, we need to hold physicians accountable for retaining people in care. Absolutely.

MEMBER OSBORN: So if a patient decides that they don't want care or that they don't want to continue care how is that --

MEMBER BRADY: It's not going to be 100 percent is what I would say. But also you know, it's a physician's responsibility to try and bring that person back to care and not just say oh well, they didn't come back. So what, you know.

MEMBER OSBORN: But is there a difference between trying to bring somebody back for care and being held accountable for the patient's decision not to return? That's what I'm trying to -- and I don't take care of these patients. I'm just trying to understand.

MEMBER BRADY: I understand but --

MEMBER OSBORN: This is a quality measure that we're holding the whole country accountable for.

MEMBER BRADY: No, I understand but I think part of the reason patients don't come back is related to their, you know, maybe the way their physician treats them. You know, there are things that it's partly the physician's responsibility that someone doesn't come back. You have failed as a clinician.

CHAIR SEPTIMUS: Michael?

MEMBER FARBER: I just wanted to make a comment on that issue of making appointments because I've been involved in that in managed care. So that yes, a physician can't be responsible always for everybody. People are -- some of them are homeless, some of them have severe mental illness and psychosocial issues.

But the issue would be, which is not being addressed in this measure, is that

is there due diligence to try to get them back. And due diligence can be in phone calls, messages, you know, by mail, even home visits. So in other words that's what's missing here is due diligence, you know, because there are situations of which the provider is absolutely not responsible.

CHAIR SEPTIMUS: Let's just keep this on track. It's about high impact, addressing a specific national healthcare goal, priority or data demonstrating a high-impact aspect of healthcare. So numbers affected, so forth. Tom?

MEMBER GIORDANO: Just to reiterate that, that we're on impact here. And this discussion is important about retention in care but this measure actually says among those who have at least two visits. So I think that discussion is important but not related to the impact of the measure.

CHAIR SEPTIMUS: Any other comments? Let's go to a vote on impact at this

point.

MS. KAHN: Okay, voting on high impact. You can go ahead and start. You have 14 high, 5 moderate, zero low and zero insufficient.

CHAIR SEPTIMUS: Okay. Let's move right to the evidence which should be very parallel to what we discussed in the previous measure. Kathleen?

MEMBER BRADY: Yes, I don't have anything to add.

CHAIR SEPTIMUS: Anyone else have anything to add? A lot of similarities. Okay, so let's vote on the evidence.

MS. KAHN: Voting on 18 evidence. You can go ahead and start. We have 17 for yes, the body of evidence meets the guidance, 2 for no, the evidence does not meet the guidance and zero for insufficient information.

CHAIR SEPTIMUS: Okay. Then let's move to opportunity and gaps and any

disparities.

MEMBER BRADY: Okay, so the data for this comes from 2009 and 2010 CMS PQRS data. The 2011 data has been requested. I don't know if we have any updates on that. The average performance rate per eligible professional was 90.3 percent in 2009 and 97.2 percent in 2010. But that's based on a small number of providers, 60 in 2009 and 61 in 2010. And so they report that data from HIVQUAL, there were 202 facilities that reported this measure in 2009 covering 9,153 patients. The facility means were 75.2 percent and 64.2 percent respectively.

CHAIR SEPTIMUS: Any discussion on that point? Aaron?

MEMBER MILSTONE: Just a quick question about the second, the HIVQUAL was it? Is that in data or is that data in a patient population that is retained in care or that has two visits, or is that in all patients with a diagnosis of HIV?

DR. CHEEVER: That's a good question. Yes, I apologize. It's not the exact same measure. It doesn't have the two-visit requirement in the denominator. It's a similar HIVQUAL measure, it's not the actual measure.

MEMBER BRADY: Okay. So I think that's actually important because based on the data that you submitted there's really not a huge gap and that there is a gap when you eliminate the visit requirement.

MR. REHM: Yes, if I can qualify that. PQRS in some ways is a self-selecting reporting system. You choose to report on the measures that you choose to report on. My guess is those who believe they have pretty good HIV care will report on that measure selectively. So you have to have a certain caveat. And we're relying on that CMS data and that's what they have available. So, we would expect -- the requested data we would expect to see higher numbers are participating

with a broader range if you will of profiles if you will of physicians who will be reporting that data. But CMS hasn't released that data yet so our hands are kind of tied.

CHAIR SEPTIMUS: Doug?

MEMBER CAMPOS-OUTCALT: This may not be the right time to bring this up but I'm really kind of confused. Because the last measure we looked at we were looking at antiretroviral therapy for everybody and we were told that it applied across the board and that the above 500 was an exception and it was a small percentage. And therefore didn't really affect the measure that much. Now, this measure applies to everybody 500 and below. So it appears to me that the last measure really only applies to people above 500 of which we had not very much evidence. Because this one is applying to -- they should be on stronger antiretroviral therapy if they're under 500. So how does the last measure differ and why am I thinking incorrectly here?

MEMBER BRADY: It's not just a difference in the numerator, there's a difference in the denominator where in the last measure it was one medical visit in a 12-month period where this one is two medical visits at least 60 days apart. So you have to meet the medical visit requirement. So people who have a CD4 count less than 500 but only have one visit in this -- in a year will not be evaluated. They won't be in the denominator. Does that make sense?

MEMBER CAMPOS-OUTCALT: I have to ponder that a little bit.

MEMBER BRADY: Yes. And so I think when we get to some of the medical visit information you will see that there's a large proportion of people with HIV who only get one medical visit in a year. And based on the guidelines in terms of following people if people are stable on antiretroviral therapy then they, you know, we talked about this before. They only need to be monitored every

6 to 12 months. So, they wouldn't be included in this measure even though they're being appropriately treated.

MEMBER CAMPOS-OUTCALT: But the recommendation for them was still to be on the stronger therapy.

MEMBER BRADY: The recommendation would be if they're stable, on therapy, they have a CD4 count of 400, you know, for a long time the recommendation would be that they should be on therapy. But they would not meet this measure because they don't get two medical visits.

MEMBER CAMPOS-OUTCALT: Right, so I go back to my point which is if the recommendation is that if you're under 500 you go on the stronger therapy the last measure will really only apply to people above 500.

MEMBER BRADY: They would be included in the last measure because it's everyone but they would not be included in this one. They would not be in the denominator for

this measure or the numerator. There's definitely overlap but the difference in the numerator is that this is only less than 500 and the other one is everyone. So, those less than 500 that are included in this one would be included in the last one, but the additional 3 percent of people who have a CD4 count over 500 are included in the last one. But in this one the difference in the denominator is the number of visits that you must have to be.

CHAIR SEPTIMUS: Did you want to say something?

MEMBER SPACH: Just real quickly. Just to clarify we're not talking about stronger therapy, we're talking about across the board therapies would be similar. We're talking about whether or not to receive therapy at all. There's no stronger therapy that we're recommending for lower CD4 count.

MEMBER MILSTONE: I just wanted to clarify. So, can any of the clinicians add or people that do this, is there -- do people

think there are, there is a gap and we heard there's no data on is there a gap in this population people that receive two visits from the data presented. Do people feel like there is a gap that we should be addressing? Because this is going to be -- it seems like harder data to capture. It's not just do you have HIV, did you get a drug, but do you have HIV, did you get multiple visits. So, I think it's -- the burden of collection would be important if there's no gap to try to fix.

CHAIR SEPTIMUS: David.

MEMBER SPACH: The Irene Hall data suggests there is a gap because they actually analyzed it and basically said for all people living in this country who have HIV only about 21 percent have suppressed levels of HIV and about 30 percent or so are actually receiving antiretroviral therapy. They did the analysis for people who were engaged in care and found that there was a gap among those engaged in care and who were receiving antiretroviral

therapy. I can't quote you the exact percentage but that Irene Hall data is available. So there is a significant percentage of people who are engaged in care and retained in care is actually the language I think they use who do not receive antiretroviral therapy.

CHAIR BROTMAN: Tom, did you want to make a statement? No?

MEMBER BRADY: I can follow up with that because I'm the PI for MMP in Philadelphia. And that data analysis does not account for the number of visits. So it's if you were seen once during actually a 4-month period you're included in that analysis. So it does not distinguish -- you don't have to have two medical visits. So, that data, you know, we don't know from that data whether there is a gap in people who have at least two medical visits at least 60 days.

CHAIR BROTMAN: Doug, I think I'm going to let you have the last word. You're

done? Okay. Anybody else? That's it. Okay, let's go to a vote on the performance gap at this point.

MS. KAHN: We're voting on 1(b) performance gap. You can go ahead and start. We have 3 high, 10 moderate, 2 low and 4 insufficient.

CHAIR BROTMAN: Okay, so that passes. Let's move on. Reliability. Place your microphone on, please.

MEMBER BRADY: Oh, thank you. So, I'm just looking at the notes that we had. From our work group it was unclear how well potent is defined and it's unclear how this would perform using EMRs outside of the test set. And there was no disparity data noted. And so that's --

CHAIR BROTMAN: Did the developers want to comment on reliability in this issue?

MS. WILLIAMS-BADER: Well, I can comment on the use of the potent ART definition. We, as I think I mentioned while HRSA was

reviewing or discussing their measure the treatment guidelines do change quite frequently for this -- treatment for HIV. So rather than have a list of drugs that would quickly get outdated we actually refer providers who are reporting on this measure to the treatment guidelines so that they can identify potent ART.

As far as the testing in the EMR and how that would perform in EMRs, other EMRs besides the test site I don't know that. I can't comment. Perhaps someone from the AMA can comment since they led the testing for the measures.

CHAIR BROTMAN: Okay. Tom? I'm sorry, go ahead.

MEMBER BRADY: I was going to follow up with some additional information. The -- what was submitted, actually there's heavy reliance on use of the CPT-2 codes which I think is problematic.

The reliability and validity data

came from the Midwest and it was again four sites. It consisted of 342 patient encounters with a visual inspection of medical records performed in 2009.

And in terms of the results automated calculation of performance was 96.6 percent, manual calculation of performance was 100 percent with a 3 percent difference.

CHAIR BROTMAN: Okay. Tom, did you have a point?

MEMBER GIORDANO: Yes. This is I guess addressing both reliability and to some extent validity. So there's the issue of the CPT codes. Absent those, and maybe that's not fair. I guess I don't understand exactly the role of those, but absent CPT codes it's extremely difficult to figure out who has a history of an AIDS defining condition because there aren't good ICD-9 codes for many of those conditions.

And it's -- the one strength of this measure is it's positioned where the evidence

is, CD4 less than 500, history of an AIDS defining illness. That's where the evidence is that you need potent ART. But trying to figure out who that is is difficult because as I said there's no good ICD-9 codes for a lot of the AIDS defining conditions and you've got to -- it's not just CD4 now, it's CD4 less than 500 ever. And so that I think presents a big reliability and validity challenge because you don't -- you need all their CD4 results. Their current CD4 could be 1,000 but they could have had a CD4 of 10, 10 years ago. And how you figure that out to me is a challenge. And whether you get the same result if you used an electronic method versus a review of paper records, et cetera, I think is an important consideration.

CHAIR SEPTIMUS: And some of that speaks to validity so let's just speak to reliability right now if we can.

MEMBER GIORDANO: I guess, I think it's reliability as well because you've got

to get the same result twice. And so if you do it electronically you get a different result than if you used paper records going back to the beginning of time.

CHAIR SEPTIMUS: Aaron, did you want to address it? Okay. Curtis?

MEMBER COLLINS: You know, this might not be the appropriate question for this discussion but more of a question for NQF. Given the similarities between this and the other measure has there been discussion about harmonizing these two? You know, I think this measure is a little bit more evidence sound compared to the last, but you know, is that a consideration or has that been discussed?

MS. BURSTIN: So the NCQA measure is an existing measure. The HRSA measure was a new measure. You'll get to hopefully the harmonization discussion and one of the things we'll ask the developers to do is in fact try to go off and see if there's a way to harmonize these.

Ideally we don't want two of these even with the nuances there. I think it actually just adds to the cacophony out there if they're slightly different.

MR. REHM: Yes, and just to add that prior to us restarting our review of the -- our existing measure set we did have several calls with HRSA and Laura and Marlene, and also included HRSA on our expert panel. So that was in the spirit of pre-harmonization.

CHAIR BROTMAN: A preview of things to come. Let's vote on reliability.

MS. KAHN: Voting on 2(a) reliability. Go ahead and start. You have 1 high, 13 moderate, 3 low and 2 insufficient evidence.

CHAIR BROTMAN: So that passes. Let's talk about validity for a minute if there's anything to add. Aaron?

MEMBER MILSTONE: So I'm a little unclear because before we talked about how these measures that relied on CPT codes were

going to be taken out and we were going to use them as e-measures, is that correct?

MS. WINKLER: Yes. We're assuming that's for all of the measures.

MEMBER MILSTONE: Thank you. So if the developers then can clarify how using an electronic query you're going to identify potent antiretroviral therapy.

MS. WILLIAMS-BADER: This is difficult. I'm not sure I can exactly speak to this on the spot. I could ask the testing team to see if they know right now how it was done. Keri said they followed the specifications. So.

MEMBER MILSTONE: We don't have those.

MS. WILLIAMS-BADER: Right, right, and I'm saying we don't have that either so it's hard for me to speak to it right now. I think we have thought about this and think that one approach we might take is doing the same thing that HRSA's doing which is actually

just to look for any combination that is not not recommended, contraindicated, rather than actually try to code for all the possible combinations of potent ART.

MEMBER BRADY: And I was going to say, and what about looking for the history of an AIDS diagnosis or a history of a CD4 count less than 500 that could have occurred many, many years ago?

MS. WILLIAMS-BADER: Right. The CD4 count we would just, we would look for the CD4 count. It wouldn't necessarily I think have to be a result that's recently been given as long as they do have access to that somewhere in the EHR as a history of a CD4 count less than.

And for the AIDS defining conditions I believe we would be able to, even if there aren't ICD-9 codes there would be SNOMED codes for these so we would actually use SNOMED as the vocabulary for those. That's actually what's recommended for -- that's the

final recommendation actually for the vocabulary you would use for diagnoses and conditions.

CHAIR BROTMAN: Aaron go ahead.

MEMBER BRADY: Can you explain what that is?

MS. WILLIAMS-BADER: We have vocabulary experts in the room so perhaps -- I don't know, Marjorie, I'm sorry to put you on the spot.

DR. RALLINS: So I think the concern earlier was if ICD-9 cannot capture some of the diagnoses and that's why our measures have been developed most recently using the clinical vocabularies that have been recommended by the HIT standards committee of the Office of the National Coordinator (ONC). So many of the e-measures that you have in front of you have been specified in accordance with those recommendations. And SNOMED and other clinical vocabularies actually tease out or do not lump diagnoses, procedures, any kind

of item that you would want to identify into one code. They actually simplify the information.

CHAIR BROTMAN: So let me see if I can -- so we don't have the e-specs and therefore they haven't been tested. I just want to make sure I understand that. Adam?

MEMBER THOMPSON: Yes. One thing I just wanted to bring up was regarding the question about finding old CD4 counts. We just had to try to do this for ADAP on the waiting list to try and prove like who had certain CD4 counts in order to qualify them to get the medication. And it was really difficult, really hard. And in fact people who came from the South which we know to have high incidence and high impact, those medical records, some of them particularly along the crescent were completely lost in the hurricane. There will be no documentation nor can you ever get it. So it was something we faced as a really big challenge and I would say it's a huge issue.

MS. WINKLER: Ed, let me just respond to your question about this. We don't have the e-specifications. We're expecting to get them.

One of the things that we do internally at NQF is we have our HIT folks take a look at the e-specs versus the written specs and to see if there's a match. And if there is then we feel that the e-specs do reflect them. So you're seeing what will be included in them once we've done that review.

And in terms of the testing, the EHR testing is what's presented here in the reliability and validity section.

CHAIR BROTMAN: All right. Aaron, do you want to have one last word? No, okay. Okay, Tiffany. I'm sorry.

MEMBER OSBORN: I just want to make sure I understand. I had a little difficult time understanding what you just said. You said that what we're looking at in front of us will be what it is once -- can you repeat

that? I didn't understand what you said.

MS. WINKLER: Once we get the e-specs in the format you saw yesterday for the hep C measures we just do a crosswalk comparison of that with the written specs that you see in the specifications sections and be sure that they both reflect the same thing. Look at the ones from the hep C measure from yesterday. That's what they're going to look like.

MS. BURSTIN: The e-specs will be based on a whole series of whatever the current standards are that are recommended by the HIT standards committee which are primarily SNOMED, AHRQ, Norm, et cetera. So we'll get those to the committee ASAP. And I think again -- and we'll have the HIT team review those. If you have issues with those we'll reassess the measure.

MEMBER OSBORN: So, but for right now we're supposed to assess based on what we're looking at in front of us, right?

MS. BURSTIN: Well, you have the testing results in front of you based on EHRs. So that, and that was using the EHR specs. So you are looking at e-measure testing based on a set of these specs. Unfortunately they were not submitted to you for review at this time.

CHAIR BROTMAN: Aaron?

MEMBER MILSTONE: Just to clarify those e-measures, those e-specs could be based on CPT codes.

MS. BURSTIN: They might include CPT but not CPT-2 which is a special data collection strategy for physician attestation. Those are different. CPT just may be the kind of -- actually Marjorie.

MEMBER MILSTONE: And then there was also --

CHAIR BROTMAN: Clarification. Hold on.

DR. RALLINS: So the e-specifications with respect to procedures

use SNOMED codes to represent procedures but also include CPT codes as well. The CPT codes for e-specifications are considered transition vocabularies because they again don't capture -- I think I heard a conversation yesterday. CPT codes tend to lump procedures into one code. So while the e-specifications may include CPT codes they are not considered ideal in actually capturing the data.

MEMBER MILSTONE: So I guess assuming that all harmonizes which would be great I do still have a concern about the definition of potent antiretroviral therapy in relation to the patient. I think it's great that you're considering revising that to match the previous measure that looked at any -- or what the -- the drugs that shouldn't be used that's published in the table and the CDC guidelines. So just guidance from NQF. We're voting now though -- are we voting on the proposed change to this or are we voting on this as using potent antiretroviral therapy?

MS. BURSTIN: You have to vote on the measure as it is before you. If the developer comes back with a change we'll ask you to reassess, if we need to, your vote.

CHAIR BROTMAN: All right. With that I think we should go ahead and vote for validity at this point.

MS. KAHN: Voting on 2(b) validity. You can go ahead and start. We have eight moderate, six low and five insufficient, zero high.

CHAIR BROTMAN: So that failed. Then we stop at this point and we're going to move onto the next measure.

CHAIR SEPTIMUS: Okay, let's keep going. We've got lots to go. Was it Robert Frost, lots to go before we sleep? Something like that. Anyway.

The next one is 0408, "HIV RNA Control After 6 Months" -- did I miss one? I meant 0407. But I said the right one. Do I get partial credit? "Six Months of Potent

Antiretroviral Therapy." This is also an NCQA so if Jenna or Bob have any comments and then we'll turn it over to Tom.

MS. WILLIAMS-BADER: Yes, just really briefly. This measure really builds off the measure that you just discussed because it does have patients 13 years and older with a diagnosis of HIV/AIDS with the two visits during the measurement year at least 90 days apart who are receiving potent antiretroviral therapy and who have a viral load less than 200 copies after at least 6 months of potent ART.

CHAIR BROTMAN: Okay. Let's discuss the measure and then go to impact.

MEMBER GIORDANO: So yes, just to go over some of the preliminaries on the measure. It's HIV RNA control after 6 months of potent antiretroviral therapy. It's from the NCQA. As mentioned it's patients aged 13 or older with a diagnosis of HIV/AIDS who had at least two medical visits during the

measurement year with at least 90 days between them who are receiving potent ART who have a viral load less than 200 copies per mil after at least 6 months of that potent ART, or of potent ART.

And that's described as any -- potent ART is described as any ART that has demonstrated optimal efficacy in results in durable suppression of HIV as shown by prior clinical trials.

This is a maintenance review and it's an outcome measure. So, I think that's the general summary. So moving onto impact?

CHAIR BROTMAN: Yes, let's go to impact, please.

MEMBER GIORDANO: Clearly HIV is common enough, 1.2 million in the U.S. and it's a leading cause of death in certain populations in the U.S. especially some minority age groups. There are a number of new infections each year. I think we know all this. And HIV RNA plasma levels assess the efficacy of ART.

RNA less than 50 is regarded as the optimal outcome although 200 copies is often used in clinical trials group, the primary clinical trials group, the AIDS clinical trials group.

RNA's level should be measured on all patients at baseline and thereafter, especially people on treatment to monitor response and to prevent disease progression. And for most individuals who are adherent to their ART and who do not have resistance viral suppression is generally achieved in 12 to 24 weeks although it could take longer in some patients.

There are a lot of studies to support high impact, that HIV suppression is good for the patient. They've cited a number of randomized trials and observational data as well as the treatment guidelines. So you know, without getting into details on that I think overall this is clearly supported by the evidence.

CHAIR BROTMAN: Okay. Aaron, did

you want to add something? No? Okay. If there's no discussion let's vote on the impact, high impact.

MS. KAHN: Voting on high impact. You can go ahead and start. Everyone press it again. There should be 18.

CHAIR BROTMAN: Push your buttons.

MS. KAHN: All right. So you have 17 for high, 1 moderate, zero low and zero insufficient.

CHAIR BROTMAN: Okay. That passes. Let's talk about the evidence for this measure.

MEMBER GIORDANO: So I guess I kind of got into that a second ago. There's very strong evidence that suppression is good. The DHHS guidelines rate achieving viral suppression as the goal of therapy and that's an A1 level rating. There are 10,000 patients summarized in those guidelines from 33 studies and so there's clearly a large evidence base to support a viral suppression.

CHAIR BROTMAN: Pretty similar amount of evidence that we've talked about. Any other discussion needed? All right, let's vote on evidence.

MS. KAHN: Voting on 18 evidence. Go ahead and start. You have 17 for yes, the body of evidence meets the guidance, 1 for no, the evidence does not meet the guidance, and zero for insufficient information.

CHAIR BROTMAN: Great, so that passes. Let's go to performance gap and disparities.

MEMBER GIORDANO: So on the performance gap the developer submitted PQRS data from 2009 and 2010 showing that in both years roughly 76 percent of persons met the standard. That was approximately 70 providers and 600 to 700 patients each year. There were no disparities data submitted as part of the application.

CHAIR BROTMAN: Can we assume that with all these measures even if there wasn't

disparity included that we've certainly heard enough comments that there really is a disparity?

MEMBER GIORDANO: I think, I mean there's clearly evidence that -- when it comes to viral suppression that there is a disparity in outcomes by -- for many demographic groups, not just race/ethnicity.

CHAIR BROTMAN: So let's go to a vote on performance gap then.

MS. KAHN: Voting on 1(b) performance gap. Go ahead and start. We have 10 high, 7 moderate, zero low and 2 insufficient.

CHAIR BROTMAN: Great. So we go onto reliability.

MEMBER GIORDANO: Okay, so for reliability the developer submitted data from a -- well, let me back up. The numerator in this case is -- I'm sorry, the denominator is all HIV-infected persons greater than age 13 with two medical visits in the measurement year

on ART for greater than or equal to 6 months. In their application they state that "on ART" is defined by CPT category 2 code.

The numerator is persons with a viral load less than 200. And that is actually not clearly specified when that viral load has to occur, at what point it's measured. Obviously sometime in the measurement year but exactly when is not clear. And what they state is that viral load less than 200 is to be captured based on CPT category 2 code that has yet to be requested is I think the language they use.

CHAIR SEPTIMUS: Just clarification. What we said about CPT-2 codes I think apply to all of our measures. So I think we need to have that resolved.

MEMBER GIORDANO: Right, right. So then in terms of the reliability of the measure they submitted data from four sites with 410 patients. I guess that's actually more validity at this point. Is this

considered an electronic and so reliability is not a concern? Or is sort of the standard?

MS. BURSTIN: Yes, which is fine. If you guys want to combine it into a single vote that's okay too.

CHAIR BROTMAN: So if you want to present all that and then we can vote on both.

MEMBER GIORDANO: Okay. Okay. So they had 4 sites, 410 patients. They did manual extraction of the measure versus an automated extraction of the measure, or calculation of the measure. And the difference between -- the medical review came up with a result of 100 percent and the automated came up with 96.6 percent. So there was only a 3 percent difference between the two ways of measuring the indicator.

I don't think that -- it's really not clear to me if that is 100 percent of persons had viral suppression or if it's 100 percent of people could be assigned one category or the other. That's not clear to me.

CHAIR BROTMAN: Does the measure developer want to comment?

MS. WILLIAMS-BADER: I believe it's the measure rate but I can ask the testing team if that's -- oh. Can you repeat the question so that our testing team can hear?

MEMBER GIORDANO: So the -- what's stated for reliability is --

CHAIR BROTMAN: Can you move the microphone closer to you? We're not hearing your voice.

MEMBER GIORDANO: What's stated for reliability testing is medical review was compared to electronic calculation and they compared electronic health record automated reports to visual inspection of the medical record. Data analysis included percent agreement at the denominator and the numerator. The automated calculation of the performance result was 96.6 percent, the manual calculation of the performance was 100 percent and the difference between the two was 3 percent.

But I'm confused as to what that means. Does it mean that 100 percent had suppression when they did the chart review? Or that percent agreement of the denominator and numerator, what does that mean?

MS. CHRISTENSEN: Hi, it's Keri Christensen from the AMA again. We worked on the testing project.

Percent agreement is a measure of reliability and it doesn't have anything to do with the actual performance rate itself. So you could have 100 percent agreement on zero percent performance or zero percent agreement on 100 percent performance.

So agreement percentage is what your typically used to where large numbers of, or large percentage of agreement is good. Ninety-seven percent of agreement would mean that on 97 percent of cases the report and the manual abstraction would agree in determining whether the patient meets the measure, does not meet the measure, is an exception. Does

that answer the question?

MEMBER GIORDANO: No, not adequately because if you've got -- if you're calculating percent agreement between the automatic and the manual you would have one agreement statistic, but you've reported an agreement statistic for automatic and one agreement statistic for manual. So I don't understand what that is.

MS. CHRISTENSEN: So the performance rate actually, the -- if you used the automated report it showed a 96 percent performance rate. And the manual --

MEMBER GIORDANO: Meaning what? I'm sorry to interrupt but what does that 96 percent performance.

MS. CHRISTENSEN: Ninety-six percent of the patients, 96 point -- I can't read it, I'm sorry -- 6 percent of the patients met the measure. The manual calculation, the manual abstraction, showed that 100 percent of patients met the measure. So the difference

is 3.4 percent between the two methods of calculating the measure.

MEMBER GIORDANO: Okay. That's clearer, thank you. So in this sample 100 percent of patients actually were suppressed.

CHAIR BROTMAN: Kathleen?

MEMBER BRADY: So my immediate comment to that, that's difficult to believe but I actually want to go back to the denominator which actually includes that persons are prescribed potent antiretroviral therapy and it goes back to the same issues for the last indicator, what's the definition of potent antiretroviral therapy.

MS. WILLIAMS-BADER: Again without having the testing specifications right in front of me I don't know exactly how they did it for the testing specifications but we are open to aligning with the definition that HRSA is going to use for their measure.

CHAIR BROTMAN: Okay. If there's no more discussion let's vote on reliability

and then validity.

MEMBER GIORDANO: I'm sorry, one more point. Could the developer specify or indicate which viral load they used? Is it any viral load that's less than 200 in the measurement year or the last viral load in the measurement year? That's not specified.

MS. WILLIAMS-BADER: Yes, that's a question that came up during the work group. I don't have the answer with me but we could clarify that.

CHAIR SEPTIMUS: So are we voting together reliability and validity together? Or individually or together?

MS. BURSTIN: Let's do it together.

CHAIR SEPTIMUS: Okay. So whenever we vote here is for both. That's what the boss said.

MS. BURSTIN: Just do validity. It's fine.

CHAIR SEPTIMUS: Do one at a time, I don't care.

MS. BURSTIN: I think we've just indicated at least our current policy is such that reliability of the data element level is not required for EHRs. What you're really looking at here is again you can argue whether this is reliability or validity. I'm with you, Peter. But for today's argument this is what we would count as validity and we'll explain it in the report. So just trying to keep it moving for you guys, but.

CHAIR SEPTIMUS: Adam, you had a question?

MEMBER THOMPSON: Yes, I just have a clarifying question and it's a follow-up to Tom's question. Because again it's this 100 percent viral suppression thing. And I just wanted to ask is it 100 percent of the patients had an indication that their viral load was monitored? Because to me that makes sense versus saying 100 percent of the patients had a viral load suppression. And to me that would affect how I would rate the reliability and

validity because I just don't think 100 percent of these patients achieve viral load suppression given the national data on this.

CHAIR SEPTIMUS: Developer?

MS. WILLIAMS-BADER: So the measure when it was tested in 2009, we have made updates to this measure based on expert feedback recently. The measure as tested in 2009 used to allow for a plan of care for patients that were not in control. So that might be speaking to -- that might help explain why the rate is so high. Again, when we reviewed this with our experts in 2012 they very strongly supported removing the plan of care component.

CHAIR SEPTIMUS: Makes a stronger measure. Okay, let's vote on validity.

MS. KAHN: Okay, voting on 2(b) validity. You can go ahead and start. I have zero high, eight moderate, seven low, and four insufficient.

CHAIR SEPTIMUS: Well, this fails.

So thank you. Do we get the chance to see you again? All right. We're going to do one more but we're going to have -- we're going to ask for any public comment. Then we're going to have a working lunch. So let's keep on track. So, Jeff, you still there?

MEMBER BEAL: Yes, I am. Thanks.

CHAIR SEPTIMUS: Okay. So this is 2082 "HIV Viral Load Suppression." This is a HRSA. So we're waiting for the HRSA developers to come up. They're coming and they'll make some comments, Jeff, and then we'll turn it over to you to lead the discussion.

MEMBER BEAL: Thank you.

CHAIR SEPTIMUS: Developers. Which one? Marlene?

MS. MATOSKY: Yes, I'll go. We're back. So we just have very brief comments about this measure that we're going to present. It's 2082, affectionately known as "HIV Viral Load Suppression."

This measure as I had briefly mentioned before has been endorsed by Dr. Sebelius for use in all HHS-funded HIV programs. Similarly to the antiretroviral therapy measure that we presented earlier we don't expect performance to be at 100 percent for this measure either. And we feel as though it has broad applicability in that it could be utilized both at the clinic level but then also at a jurisdictional level, you know, a metropolitan area, a city, a state and even nationally. So those are the only comments that we have.

CHAIR SEPTIMUS: Okay. Jeff, let's start off with impact, please.

MEMBER BEAL: All right, thanks. The measure description is the percentage of patients regardless of age with a diagnosis of HIV with an HIV viral load less than 200 copies at the last viral load test during the measurement year.

The numerator is the number of

patients in the denominator with an HIV viral load less than 200 at last HIV viral load test in the measurement year. And the denominator is the number of HIV patients regardless of age with at least one medical visit in the measurement year. There were no patient exclusions.

For impact our work group was unanimous in rating as high. It was supported by clinical trial evidence of antiretroviral therapy reducing HIV-associated morbidity and mortality as well as antiretroviral therapy improving quality of life. And the emerging evidence of earlier antiretroviral therapy decreasing HIV-associated complications. Antiretroviral therapy has also been shown to reduce transmission.

There was discussion in our group about the data being the strongest for the adolescent and adult population with less support in the data for the pediatric population. And there were comments about

less support of ARV therapy at the higher viral load levels as we've heard before.

CHAIR SEPTIMUS: Thank you, Jeff. So let's -- I think this is reasonably straightforward in terms of impact but I want to make sure. It looks like Mohamad is -- he wants to speak.

MEMBER FAKIH: You know, my question, why is it the last viral load, not any of the viral loads within that year? Because the issue is compliance of patients. You may do everything you can do for the patient, that patient may not become compliant. As healthcare providers if we show that we reached that level, you know, for me it's a positive thing about the work of the healthcare worker. Just an idea.

CHAIR SEPTIMUS: Marlene, do you want to comment on that?

MS. MATOSKY: So, I think it's two fold. And first I would say that we wanted to choose the last viral load in that we wanted

the most current information that was most -- even though these measures are a snapshot in time we wanted to be using the most current information for populating this measure.

And then also I think is though when you think about measure feasibility and usability we wanted to have something that was very straightforward and easy to calculate. I mean, we could have chosen the lowest viral load you know ever in the measurement year, the first one, the last one, so we chose something that we felt as though was the most readily available and most feasible.

CHAIR SEPTIMUS: Thank you. Any other comments? Then let's vote on impact.

MS. KAHN: Voting on 1(a) high impact. You can go ahead and start. Eighteen high and one moderate, zero low, zero insufficient.

CHAIR SEPTIMUS: Okay. Jeff, we're going to go to the evidence.

MEMBER BEAL: The evidence was

clinical practice guidelines specifically referencing the DHHS guidelines whose treatment recommendations are based on the analysis of six randomized controlled trials. One of those is a meta-analysis of nine randomized controlled trials. In addition, there were eight observational studies.

The quality of the randomized trials was high and observational studies were large in size. Our group rated the evidence as moderate to high with comments made about the data for starting ARV therapy greater than 500 and comments regarding the smaller body of evidence present to support the recommendation of treatment as a means of reducing transmission.

CHAIR SEPTIMUS: Okay. Any comments on the evidence? Seeing none we'll vote on the evidence.

MS. KAHN: Voting on 18 evidence. Go ahead and start. You have 18 for yes, the body of evidence meets the guidance, 1 for no,

the evidence does not meet the guidance and zero for insufficient information.

CHAIR SEPTIMUS: Okay. The next is opportunity and gap. Jeff?

MEMBER BEAL: The majority of our group felt the measure could identify areas of improvement for clinicians in its monitoring as was supported by data from the Medical Monitoring Project showing 77 percent achieved viral load suppression at most recent test, additional data from King County showing 65 percent achieved undetectable at last test and data from Kaiser Permanente showing that 94.5 percent achieved undetectable at last viral load if they were known to be on ARV therapy with 69 percent achieving undetectable when looking at all HIV-infected populations in their data set.

Disparities were identified in viral load suppression by race as well as by age and sex.

CHAIR SEPTIMUS: So just a point

of clarification. We're looking at undetectable in less than 200.

MEMBER BEAL: Yes.

CHAIR SEPTIMUS: And I would ask people who perhaps do this every day is that actually the new standard. Versus --

MEMBER BEAL: It's the definition in the DHHS guidelines, yes.

CHAIR SEPTIMUS: I should have raised this before but obviously things have changed and didn't know if that should necessarily affect our decision on this particular measure but I think it has changed. Tom?

MEMBER GIORDANO: So I would say that the goal is still an undetectable viral load, maximal suppression, which most assays now it's less than 50, less than 48, less than 20.

However, blips in viral load that are thought to probably not be clinically relevant, at least immediately clinically

relevant, are not uncommon. And so what is recommended is you don't consider a regimen to have failed until you have reproducible viral loads over 200.

The empiric data to back that up are, you know, that 200 is the right cut point are not -- there's not a ton of them. However, I think that most experts would agree that that's a reasonable standard and that's only a minor component of this measure. So I think it makes sense.

CHAIR SEPTIMUS: I just want to raise a point of discussion just to know that there are different standards. And obviously Tom's right, some people will occasionally get above that magic number and then the next time you test them they're fully suppressed again. So I just wanted to bring that up just as a point of discussion.

DR. CHEEVER: I just wanted to add on that the reason it's less than 200 on the adult guidelines is because there's work by

Dr. Silicano at Hopkins that shows that those blips that people -- that do occur often are related to what they think is just release of virus from already-infected cells and not breakthrough of antiretroviral therapy.

CHAIR SEPTIMUS: Any other thing about the gap? Let's vote.

MS. KAHN: Voting on 1(b) performance gap. You can go ahead and start. We have 7 high, 12 moderate, zero low and zero insufficient.

CHAIR SEPTIMUS: Okay. Now we go onto my two favorite elements, reliability and validity. Jeff?

MEMBER BEAL: All right, reliability and validity were assessed only at the measure level. Reliability testing was done through the multi-site HIV Research Network which is inclusive of community and academic HIV providers in four major geographic regions in the United States. Nine out of the eighteen sites which used ultra-sensitive

viral load testing were included in the reliability analysis with patients included in the analysis if they had at least one visit in a 12-month period. The group, our group, work group majority assessed the reliability as moderate noting good sampling and well-defined testing data.

CHAIR SEPTIMUS: Any comments from either the work group or the committee? Seeing none we will vote on reliability.

MS. KAHN: Voting on 2(a) reliability. You can go ahead and start. Can we have everyone press it one more time? You have 2 high, 17 moderate, zero low and zero insufficient.

CHAIR SEPTIMUS: Okay. Next we're going to go to validity.

MEMBER BEAL: All right. The analysis of this data was by face validity established through a technical work group of leading researchers and physicians in HIV retention, care and treatment as well as

governmental and non-governmental public health officials across the country.

Experts in the work group presented the most current research to the group and it was noted that often the principal investigator of the study made the presentation. The group discussed and identified the data elements with a simple majority defining consensus on the final set of measures.

Additional validity was then gained through structured webinar presentations with national representation of Ryan White providers who were asked to implement the measures into their quality management program and to provide feedback which was gathered at a later webinar. On review our group had assessed the validity to be moderate.

CHAIR SEPTIMUS: Comments from the committee? Then we'll vote.

MS. KAHN: Voting on 2(b) validity. You can go ahead and start.

CHAIR SEPTIMUS: I guess we're

going too fast for you.

MEMBER GIORDANO: Ed or Steven? I'm listed as one of the people in that expert panel on this. So I think I want to abstain from voting on this particular issue.

CHAIR SEPTIMUS: It's up to you.

MEMBER GIORDANO: I'll abstain.

MS. KAHN: Two more. One more time. We have 1 high, 17 moderate, zero low and zero insufficient.

CHAIR SEPTIMUS: Okay. So either you're getting tired or you're getting hungry. Both? Okay, usability.

MEMBER BEAL: The data presented discussed the usefulness of this measure to providers of HIV care and treatment. And this measure is currently used by National Quality Improvement Project focused on retention in medical care.

The Centers for Medicare and Medicaid has endorsed this measure and Ryan White providers using the measure report this

measure as being meaningful and useful for quality improvement activities.

DHHS, the Veterans Association, Kaiser Permanente and HIVMA have endorsed this measure. Our group majority was high to moderate.

CHAIR SEPTIMUS: Comments? Then let's vote on usability, please.

MS. KAHN: Voting on usability. You can go ahead and start. We have 10 high, 9 moderate, zero low and zero insufficient.

CHAIR SEPTIMUS: Okay. Feasibility, please.

MEMBER BEAL: The clinical data of the HIV viral load are generated, tracked and monitored as a routine of patient care. The data points are available in electronic health records and from lab reports and there were no identified inaccuracies or unintended consequences of measurement identified during testing. Our work group rated feasibility as high to moderate.

CHAIR SEPTIMUS: Comments? You guys are ready to vote before comments now. This is -- no comments, we'll vote.

MS. KAHN: Okay. Voting on feasibility. You can go ahead and start. We have 8 high, 11 moderate, zero low and zero insufficient.

CHAIR SEPTIMUS: Okay. Then the last in this set is of course the -- whether this is applicable measure to be endorsed. Does it meet the criteria.

MS. KAHN: Does the measure meet NQF criteria, yes or no. You can go ahead and vote. We're one short. There we go. Eighteen yes and one no.

CHAIR SEPTIMUS: Excellent. This measure is finished. So before we go to lunch we're going to ask the operator if there's any public comments.

OPERATOR: At this time I'd like to remind everyone in order to ask a question press * then the number 1 on your telephone

keypad.

CHAIR SEPTIMUS: And anyone in the room who would also like to make a comment please let me know.

OPERATOR: At this time there are no questions.

CHAIR SEPTIMUS: No questions here. Okay. So here's the plan. Lunch has arrived. We'll take about 10, maybe 15 minutes to take a bio break, get your lunch and then we'll try to reconvene before 12:30 and then work through lunch until we finish. So we'll see you back here let's say no later than 12:30.

(Whereupon, the above-entitled matter went off the record at 12:09 p.m. and resumed at 12:30 p.m.)

CHAIR SEPTIMUS: Okay, we have discussed. We have so much momentum now with these HIV measures, rather than change course now I think it would be disruptive. So we're going to continue with the next set of HIV measures starting on newly enrolled in medical

care, 2081. Michael, are you ready? Okay and this is going to again be HRSA. So we'll ask our developers to make a few comments and then we'll turn it over to Michael.

MS. MATOSKY: So as you've probably figured out we have three measures that are coming back to back to back that are all related to medical visits. And we feel as though retention in care as you've probably heard from the discussions that have occurred yesterday and today is a significant issue within the context of HIV care, treatment and prevention. And it's one of those things where it's not as straightforward as viral load suppression. We know how it impact suppression, we know how to measure it, we know what it means to be suppressed or not to be suppressed whereas retention, the body of evidence is growing, expanding rather rapidly, even as frequently as the last few months.

And so based on the best science that we had when we were developing these

measures we came at retention from a couple of different angles. And so we see these more -- we don't see as a composite measure, we see them more as a suite of measures in that they can be working together when implementing and measuring retention.

And we also pulled out some very specific aspects of retention and very specific populations when it comes to retention because we know that there is a little more evidence suggesting that there are certain populations that are more vulnerable for loss to care and in need of retention. That's it. Thank you.

CHAIR BROTMAN: Okay. Okay. And with that, Michael, I think it's your presentation. Thanks.

MEMBER FARBER: Yes, I just wanted to make also a comment or two. This measure is I think a very sentinel measure because it gets at the point of are visits necessary. And the other measures, there was something that actually had to be done at the visit,

something specific. These measures, there isn't anything.

And that was one of the weaknesses of the measure is that it doesn't define what actually occurs at the visit. But we know that the purpose of the first one of newly enrolled is that there should be visits that occur across the year. And this, the description of this measure is in the numerator of visits every 4 months over a visit in the first month. And it's for all HIV patients, all ages. And let's see. The issue with the visits is what things can be prevented. And all of the treatments that are related to HIV all will come from a visit. So the things that have been shown in the studies to try to show that this is a benefit is that there is first the increased survival, and that's because people get CD4 counts earlier. If they're abnormal they get treated with antiretroviral drugs earlier.

One of the issues also is that it's

more than just studies because this is also opportunities to counsel people and to discuss lifestyle and behavior and to give people support and to again perhaps keep them with visits so that they continue beyond the first year.

So, I guess the first discussion would be on impact. And our group felt for the most part that this was high and one moderate in its importance. And so that's the first issue.

CHAIR BROTMAN: Okay. Any comments or discussion on the impact? Let's go to a vote.

MS. KAHN: Voting on high impact. You can go ahead and start. We have 14 high, 3 moderate, 2 low and zero insufficient.

CHAIR BROTMAN: Okay. That's great. Let's go to the evidence, Michael.

MEMBER FARBER: Now as I said, the evidence which has been many studies show again as I stated earlier increased survival,

increased use of CD4 and also issues with being on antiretroviral drugs. And so I think that we said the quantity of evidence was good.

The quality, the issue again where there was some concern is that it really -- this measure doesn't define what does happen at the visit. So that this measure you know is one in which the visit is probably with an HIV specialist so that it's a highly specialized person and they do have protocols of what they're going to do in the visit. But the only measurement for this visit is that you came to it.

The consistency of the studies was felt to be pretty strong because they all, many of them looked at the same issues that I already brought up to try to define that there was a benefit that could be measured to these visits in the first year. So the group felt that this was also again high and one moderate.

CHAIR BROTMAN: Any other comments with evidence specific to this? Doug?

MEMBER CAMPOS-OUTCALT: So what would you say would be the level of evidence comparing to three visits per year to two?

MEMBER FARBER: That was the weakness and in fact I was -- that was a stress for me. And that is that we haven't really defined -- the two measures that I have been assigned don't compare it. The studies don't compare visits to a different frequency of visits. So that what it is is that studies show that visits were useful but what I think the weakness to me of this measure was is that we haven't really defined what is the optimal number. But that these three visits in the year did improve many different parameters. Whether four or five visits would have been better, but also they may have been harder to insure. So that is a weakness.

CHAIR BROTMAN: Peter and then Tom.

MEMBER HAVENS: You said you assumed that these visits were to an HIV care provider but that's not actually specified in

the --

MEMBER FARBER: Correct. But see, you got in because you were newly diagnosed so that it would seem that some referral was made to someone based on your diagnosis which might have occurred from somebody else.

MEMBER HAVENS: I don't think that's required of the guideline.

MEMBER FARBER: Right. It's not required.

MEMBER HAVENS: Right? This could be all visits to a family practitioner, all visits to a reporting obstetrician/gynecologist.

MEMBER FARBER: Correct. And many -- well see, I think that in general practice today that there are few doctors who are going to continue seeing people across the year who have no experience whatsoever with HIV. But in the first year there may not be any treatment decisions made. But there are different tests that need to be performed so I think that --

I agree that there could be a great variability in what occurs in the visits in this because we haven't specified that.

CHAIR BROTMAN: Tom?

MEMBER GIORDANO: On that issue typically what HRSA has -- I believe HRSA has this definition of a provider visit is a provider who is an antiretroviral prescriber. So someone who in that clinical setting would manage HIV. But I don't see it in this guideline anywhere. I don't know if the developer wants to comment on that.

CHAIR BROTMAN: Please.

MS. MATOSKY: Sure. So when we tested this measure we used visits that were conducted by a physician, a nurse practitioner or a physician's assistant. And in the event that this measure gets endorsed and we go to e-specification we would use the appropriate CPT codes that would be utilized by those folks I just outlined.

MEMBER HAVENS: That's not the

question. Those people could all -- those are all licensed independent practitioners. So the question is what is the quality of the LIP that you would be looking at. Would it be a specific HIV-focused provider or any licensed independent practitioner which is not specified here. And as written this could be a family practice nurse practitioner for visits in the first year with no experience in HIV care.

MEMBER FARBER: It's three visits.

MEMBER HAVENS: Well, whatever, whichever one this is could be to a non-HIV specialty care provider as written.

DR. CHEEVER: So yes, that is true, that could occur. When we've looked at studies of people that have not been linked to care and not been retained in care, if you look at actually where those missed opportunities are and where they're showing up they're not generally showing up in a primary care setting. They're showing up in an emergency room or

other settings where those missed opportunities are in terms of re-engaging them in care.

So in fact if they were seeing a -- they did stay in care with a family practitioner and saw that person over the course of the year that would count towards them being in care in this measure.

MEMBER FARBER: I think also that what the studies show is that the, you know, the absence of visits leads to poorer outcomes. So, but the results of most of the studies were really in the 60 to 70 percent range. So as far as making visits. So there is considerable room here for improvement and also for questions on defining how to improve that.

CHAIR BROTMAN: Did you want to respond?

MEMBER HAVENS: Well, actually I had a question for Tom who's been involved in some of this work. The question would be is the practitioner type involved in the visit

associated with the outcome of interest.

MEMBER GIORDANO: I think that the research base, the evidence base is in -- with visits for -- with prescribers of ART or potential prescribers of ART. I understand the developer's comment but I think if you looked at the evidence it would be mostly based in visits with someone who is able to manage HIV.

And earlier it was brought up, the number of visits that should be required, is three the right number. Should it be two, should it be four. There's some research on that as well and that question is to some extent unanswered. But I think, remember this is for patients new to the clinical setting and so I think there is -- I think three visits is probably a clinically reasonable approach.

Every time -- if you require more visits of course you're going to have a higher -- you're going to exclude more people. You're going to find more people who are not retained.

But some of that may be misclassification. The lower the number of visits required obviously you're going to get people who meet the measure but are not actually truly retained. So it's a balancing act there. And the research to date on what the right number is has found that there is no one precise number. But that shouldn't stop us from trying to improve quality.

CHAIR BROTMAN: I think the measure developer wants to comment.

DR. CHEEVER: So just one more thing in addition to my previous comment. There's -- in the U.S. we really don't have a definition of an HIV expert per se. It would be -- it's very hard to define those people. It's not like a cardiologist where you have your certification. So, although there are certifications they're not used universally, et cetera. So that's just a consideration we have that in fact if you said go back and only have this for HIV prescribers that would be

very, very hard to get that definition depending how different licensing and prescribing is done in different clinical settings. For example, where one provider is -- the unit of the prescription is just the whole organization.

MEMBER FARBER: Well, I'll say this. I can see that in some, especially rural communities that -- for this measure that a person that is not an antiretroviral prescriber could be initially the provider and could do a very adequate job of following their CD4 counts and counseling them, and then making a referral when they actually need antiretroviral therapy because that might, especially in the state I'm in, Vermont, there's only one real place that you can go to and that's Burlington. So you would have to make a referral.

CHAIR BROTMAN: Adam.

MEMBER THOMPSON: Yes. The question I have has to do with whether or not

the evidence supports the need for so many medical visits for individuals who don't necessarily have a gap in care but have just transferred their care. So if they've been retained in care over a 10-year period and they're just moving hospitals it seems like that would increase the burden on the patient to have to visit their doctor in that first year.

CHAIR BROTMAN: That is an interesting point. Aaron, did you want to say something?

MEMBER MILSTONE: I just wanted to make sure I understood population. So if I decide this year on New Year's that I'm going to go get HIV tested and I go the first week in January to my PMD to get HIV tested and it's positive then I think I fall in your population, right? It's a medical visit in the first 4 months of the year and I'm a new patient, I'm newly enrolled. And then my PMD like was mentioned by Michael says I don't treat HIV,

go to, you know, the speciality clinic and I get referred. What happens to the primary care doctor that enrolled me in the first 4 months but that I don't see after that because I get referred for specialty care?

CHAIR BROTMAN: Measure developer.

MEMBER MILSTONE: Just to follow up on that. I ask because this was validated within an HIV research network so in that population they're not referring, it's staying within house. But if it gets applied broadly you're going to have to also deal with all these other providers that do refer.

MS. MATOSKY: So in that instance that you did mention, that initial primary care physician if they were utilizing this measure, that patient would make it into the denominator but not make it into the numerator.

CHAIR BROTMAN: Go ahead, Kathleen.

MEMBER BRADY: Well, I guess I'm -- so who is this measure for? Like what's

-- you know what I mean? What's the population group that's the intended audience?

MEMBER FARBER: All HIV patients any age.

MEMBER BRADY: But I meant for the -- like this is going to be measured at the --

DR. CHEEVER: I guess we had assumed that it would be used for people that were treating HIV infection, that those are the populations that they would be studying how well they're retaining people in care that have come to them for HIV treatment. But obviously there are others like that example.

CHAIR BROTMAN: Tiffany.

MEMBER OSBORN: Just a thought for consideration. Again I don't take care of the primary issues associated with HIV but I have worked in areas such as when I worked in D.C. or when I worked in Virginia or when I worked in South Texas where there's migratory workers. So people who come in, they will get treated

and they may go to either another area of a country or another country. And I'm just wondering how that impacts this.

CHAIR BROTMAN: Measure developer, if you care to comment.

DR. CHEEVER: Yes, I think there are many cases. I think for many of us that work in urban populations there are patients that are getting incarcerated and we know they're incarcerated. So once again this is looked at in the concept of a performance measure where it wouldn't be 100 percent and could you account for those patients that you haven't seen or have fallen out of care. You're going to be at an 80 percent level and the issue is that this person was only here for one visit because they migrated. We know that they were following the bean crop up the coast or whatever the particular case was.

CHAIR BROTMAN: Mary.

MEMBER BLANK: I just also wanted some clarity to follow up to the question down

there where the example of being diagnosed on January 1st gets you into the denominator but if you're diagnosed May 1st you're not in the denominator?

MS. MATOSKY: So going back to our initial statements when we opened with the previous measures when we think about performance measurement we also think about the quality improvement piece. And just from our own perspective within the HIV/AIDS Bureau we think about performance measurement not being done once a year. We think of it going on or occurring on a rolling basis. And we know that many of our jurisdictions are implementing performance measures and they're measuring them quarterly, bi-monthly, what have you. So that if you weren't picked up in one measurement period you may be picked up in the next or the subsequent measurement period.

CHAIR BROTMAN: Aaron, did you want to follow up?

MEMBER MILSTONE: I guess I just wanted to follow up on my -- just the impact. Because we had a discussion yesterday about how the measures that we're deciding on shouldn't be good for just internal quality improvement. They should be kind of broadly applicable. So I understand that these are broadly applicable to HIV providers, but then how do we think of the measure. Is this just targeted for HIV providers? Is that broadly enough?

MS. BURSTIN: I think it's a question for the committee. I mean, HIV care is fairly specialized so I think it may be appropriate to have HIV-specialized measures. But I think the provider question you raise is a good one.

CHAIR SEPTIMUS: Well, I guess anyone caring for an HIV patient who takes that responsibility should meet a certain standard of care.

MEMBER MILSTONE: I guess I feel

like it depends on how the measure is going to be used, right? If this is going to be pay-for-performance then you don't want to ding all the primary care providers who are referring patients and saying I'm not managing this, I'm going to refer and therefore I'm not meeting my expectations. If it's going to be used for assessing quality of care within HIV providers I think it's fantastic.

So I guess we're -- I know I've been speaking to people at other times, I think we're all struggling with what the intent of the measure is. So that would be helpful.

MS. MATOSKY: So thinking about the meaningful use in the PQRS measures the eligible professionals decide which measures they're going to report on. So if I'm, you know, thinking about if I'm a cardiologist I might be more inclined to use cardiology measures versus HIV measures. So you know, thinking about those are the two probably broad programs that outside of Ryan White that these

may be utilized in.

CHAIR BROTMAN: Any other discussion? Go ahead, Doug.

MEMBER CAMPOS-OUTCALT: I think we ought to vote on the evidence question because we may not go any further. We've already heard there's no evidence to support --

CHAIR BROTMAN: I agree. Mary, is your card up for a reason? I'm sorry. Okay. So let's vote for evidence if there's no other discussion at this point.

MS. KAHN: Voting on 18 evidence. You can go ahead and start. Can we have everyone press it one more time, please? We have eight yes, the body of evidence meets the guidance, two no, the evidence does not meet the guidance and eight for insufficient information.

CHAIR SEPTIMUS: Okay, well this measure fails. Okay. Let's go to 2079, "Medical Visit Frequency" also a HRSA developer. So we'll let them make their

initial comments.

MS. MATOSKY: We don't have any additional comments.

CHAIR BROTMAN: Okay. In that case let's go to the presentation. Adam?

MEMBER THOMPSON: So this measure is looking at medical visit frequency. And the brief description of the measure is the percentage of patients regardless of age with a diagnosis of HIV who had at least one medical visit in each 6-month period of a 24-month measurement period with a minimum of 60 days between medical visits.

The difference between this one and the other ones that we're going to look at is really the measurement period which is looking at a 24-month period rather than a single year period. And it's looking at not necessarily adherence to the visit but looking at how frequently an individual made those visits over a 2-year period. And it's not specific to newly enrolled but rather any individual in

care.

The only denominator exclusion are any patients who died at any time during the 24-month measurement period and they do require that you document the date of that death. So you have to prove that that person actually is deceased.

When looking at the impact for the -- let me pull this out here. When looking at the summary of high impact it was shown that linkage to HIV medical care shortly after diagnosis and continuous care thereafter provide opportunities for risk reduction counseling, initiation of treatment and other strategies to improve health outcomes.

It showed that each no-show clinic visit conveyed a 17 percent increased risk of delayed viral load suppression which we talked earlier about. And also that the consistency of visits during the first year, having that primary care visit, there was a link between that and survival. Also, that CD4 counts were

significantly greater amongst those with optimal retention.

In our work group when we voted on this everyone agreed that it was of high impact.

CHAIR BROTMAN: Any discussion on this point? Let's go and vote for high impact at this point.

MS. KAHN: Voting on 1(a) high impact. You can go ahead and start. We have 13 high, 5 moderate, 1 low and zero insufficient.

CHAIR BROTMAN: So that passes. Adam, why don't you tell us about the evidence.

MEMBER THOMPSON: When looking at the evidence they cited a systematic literature search to produce an evidence base restricted to randomized controlled trials and observational studies that had at least one measured biological or behavioral endpoint. The recommendation that they're using focused on monitoring retention in care was based on two studies.

They also cited the Department of Health and Human Services guidelines that were based both on adults and adolescents with 14 studies examining the impact of treatment on reducing morbidity and mortality, 8 of which of those studies focused on the impact of treatment on preventing transmission and 3 of those studies that supported the frequency of CD4 count monitoring and 9 supporting the frequency of viral load monitoring.

The quality of the body of evidence was cited as two well-designed analyses of cohort studies and they had the consistency rated between the two studies showing that they were consistent in the studies that were cited.

CHAIR BROTMAN: Any comments upon the evidence related to this? All right. Seeing that there's not let's vote on the evidence.

MS. KAHN: Voting on 18 evidence. You can go ahead and start. You have 14 for yes, the body of evidence meets the guidance,

4 for no, evidence does not meet the guidance and 1 for insufficient information.

CHAIR BROTMAN: Okay. That passes. Let's go and talk about the performance gap and disparities, Adam.

MEMBER THOMPSON: When looking at the performance gap they cited data that show that individuals as they were progressing in their care over a period of time there was a reduction in their ability to maintain their medical visits, showing that there was a need to measure this.

They also cited their own internal data that looked at only 42.6 percent of the patients had met the HRSA criterion for retention to medical visits. They also have their data broken out by disparities and do identify that there are disparities in this and the data is presented a little bit later in their validity testing.

CHAIR BROTMAN: Anybody want to add anything or comments? Okay, let's vote on the

performance gap.

MS. KAHN: Voting on 1(b) performance gap. You can go ahead and start. You have 6 high, 13 moderate, zero low and zero insufficient.

CHAIR BROTMAN: Okay. It passes that. So reliability.

MEMBER THOMPSON: When looking at the reliability the data source that they used were electronic health records. I believe it was a little bit more explained a little earlier that it was used from a bunch of different data sources.

Because they tested on the electronic health record they were not necessarily required to submit reliability testing. However, they did. The sample was based on a representative sample that matched CDC incidence data that was also geographically representative. They did a signal-to-noise ratio and supplied that information showing that their test results were reliable.

CHAIR BROTMAN: Any comments?

MS. WINKLER: I just want to make a clarification. These measures are not submitted with EHR specifications so it is different. We are looking for the reliability and the validity. That's why you do have the data for both.

CHAIR BROTMAN: Tiffany.

MEMBER OSBORN: So, one question on this, and really it's to our colleagues who deal in this area. And it kind of goes back to one of the discussions that we had previously. So, this is judging physicians and hospital systems based on whether or not the patient comes in for an appointment if I'm understanding correctly. Is that -- am I understanding that correctly? Right.

So the question that I have is if you have set up a system -- I mean, it's really the issue is a system set up to try to support the patient in coming back. Because at the end of the day it's still a patient's decision

to come back or not come back, and short of going out and forcing someone to come in, you know, you can't force a person to take advantage of the medical care that you're offering to provide them. So I just wanted to --

CHAIR BROTMAN: Measure developer, do you want to comment?

MS. MATOSKY: Does anyone else want to comment before I do?

MEMBER FAKIH: If you don't mind. You know, I see this in our practice. We have a private office where the attending physicians see HIV patients and we also have a fellows' office health clinic. And we staff them, the attendings staff both. And you see the no-show rate in the residents' or fellows' clinic are much, much higher than the faculty. You know, it's the same people. So there is -- it may be disparities but there is an issue with patients, you know, patient compliance to come in or other issues with their social status.

CHAIR BROTMAN: Tom, did you want

to make a comment?

MEMBER GIORDANO: Yes, I mean clearly there's a patient factor here that's out of the clinic or the provider's control, there's no doubt about that. But I think what the measure encourages providers to do is look at what they can do to maximize retention. And that's not -- that could be co-location of services. It could be making sure you have good customer satisfaction programs. There are a lot of things that could be done. I think that what we know that if you're not in care you're not going to do well. And so this is to try I think to drive people to at least pay attention to the issue.

Everything you said is right, there is no way to completely remove the patient factor. But I think that should not -- you shouldn't get a pass on this just because the patient has decided not to come back. What are you doing to try to re-engage the patient? Are there ways to help patients stay in? So,

it does I believe reflect a quality at the provider level. That's my own.

CHAIR BROTMAN: Kathleen and then Aaron.

MEMBER BRADY: So from my personal experience at the health department in overseeing the quality management program for our Ryan White Part A grantees in the Philadelphia EMA which represents about 15,000 people engaged in HIV-related medical care. This issue comes up all the time.

And you know, in the quality improvement projects that we've seen especially around retention in care it's very easy to blame the patient. And it's a little bit, a huge pet peeve of mine in that there are plenty of things that we can do to really re-engage people in care like Tom said. And you know, we've seen from those quality improvement projects that you really can, there are things that providers can do to actually impact this measure. And it's not all on the

patient end. And it's very easy to blame the patient but I think we have to get out of the habit of doing that.

CHAIR BROTMAN: Thank you. Aaron?

MEMBER THOMPSON: One other thing I would add. This issue was raised in our work group call and the measure steward did respond that there was not the expectation that there would be 100 percent performance on this measure, that there was leeway around that for the ability for patients to not make their visits. And the expectation was that the provider would not necessarily be dinged for that.

Also, just to mention this was only tested on one of the two data levels. So the highest according to NQF criteria that we could rate it would be moderate.

CHAIR BROTMAN: Okay. Seeing there's no other cards up let's vote for reliability.

MS. KAHN: Voting on 2(a)

reliability. You can go ahead and start. You have 2 high, 13 moderate, 3 low and 1 insufficient evidence.

CHAIR BROTMAN: That obviously passes. Adam, is there anything to address validity specifically?

MEMBER THOMPSON: Just that face validity was established systematically using a modified delphi process which is one of the NQF recommended processes. They also had this as with the other measure a structured webinar around Ryan White providers. And that it was deemed, the measure was found to be important, usable and feasible by the technical work group.

The only thing was that testing was not performed any of the excluded patients so there was no threats to validity assessed.

CHAIR BROTMAN: For face validity do you have the number of -- in that? And was there a kappa score? Oh I'm sorry, you don't have a kappa score. Okay. Peter, go ahead.

MEMBER HAVENS: So again, the question of is this measuring what you really want it to measure would depend on which population of providers this is applied to. And so can I get some feedback from the developers that the intent of this is to measure retention in care in programs that predominantly serve people with HIV. Is that a true statement?

DR. CHEEVER: Yes, we envision this for people that are managing HIV infection in a group of patients with HIV.

MEMBER HAVENS: So that gets around the problem that we had with the initial, with the first measure where primary care people who don't usually do it, that would not apply in this context.

CHAIR BROTMAN: Okay. Any other comments? Let's vote on validity then.

MS. KAHN: Voting on 2(b) validity. You can go ahead and start. You have zero high, 16 moderate, 1 low and 2 insufficient

evidence.

CHAIR BROTMAN: Okay. So that passes. Let's go onto usability.

MEMBER THOMPSON: Related to usability the intended use is for public health and disease surveillance, public reporting and quality improvement with benchmarking. The current use is quality improvement with benchmarking.

The technical work group that they utilized did see a utility in this being publicly reported. They also have intentions to submit this for the EHR incentive program.

And they also believe that this measure fills a gap in measurement related to retention in care, and it's based on newer literature in the area and sort of fits a need that's not currently being measured.

CHAIR BROTMAN: Any comments on usability? Tom. Tom, put your mike on, please.

MEMBER FILE: Thanks, Adam. Very quickly, you mentioned about using for benchmarking. Have they have any idea of what that level of benchmark should be?

MEMBER THOMPSON: That I would have to ask the measure developer. They did present the data here and they have four data points that they looked at with around 146 providers reporting and each one roughly around anywhere from 62 to 64 percent is where they were.

MS. MATOSKY: So in the event that this measure gets endorsed and we get it into meaningful use and PQRS we're going to follow the methodology that ONC has suggested with setting benchmarking is that we wait until the measure is established, they've collected a reasonable amount of data and therefore after set a benchmark.

CHAIR BROTMAN: Tom.

MEMBER FILE: This goes to a point that's been made many, many, many times about the concern for inappropriately dinging people

or providers. And I've always felt, you know, that these measures here serve a purpose so that you can actually promote improvement actually. I mean, you can't improve things you don't measure. So you measure what it is and then you seek improvement. And then you know, you establish maybe what a benchmark should be, maybe 80-90 percent of whatever that measure is that actually accurately reflects what is good care.

To really expect -- well, some of these should be 100 percent that we've talked about as far as processes of care, but these types of things, I mean to expect 100 percent would be unrealistic. I mean, you mentioned that. And so I think it's just -- I just wanted to bring that out. And so I'm glad you actually are talking about assessing with a benchmark.

MS. MATOSKY: You know, interestingly enough if this comment had come up with the viral load suppression I would have had a better answer for you in that we have

a national HIV/AIDS strategy and it talks about viral load suppression among certain groups. And that document has actually set some benchmarks -- has set a benchmark for us to achieve. But we don't have -- at this point have any national benchmarks.

But as you can tell you know from the data that we've presented from the HIV Research Network and some -- an internal, or sorry, a National Quality Improvement campaign there's plenty of room for improvement but we're not at a point to say this is where we need to be by this time.

CHAIR BROTMAN: Okay. If there's no other questions let's vote on usability.

MS. KAHN: Voting on usability. You can start. Four high, twelve moderate, three low and zero insufficient.

CHAIR BROTMAN: Okay. So that passes. It's not a stop measure but we'll go onto feasibility.

MEMBER THOMPSON: For feasibility

all of the elements are contained within electronic claims. They did not list to their knowledge any known inaccuracies. And in the data collection strategy they did say that previously they had asked for persons who were incarcerated to be excluded from the denominator but in difficulty in coding that data they had eliminated that as one of the denominator exclusions.

CHAIR BROTMAN: Any comment? All right. Let's vote on feasibility.

MS. KAHN: Voting on feasibility. You can start. You have 4 high, 12 moderate, 3 low and zero insufficient.

CHAIR BROTMAN: And finally let's vote on suitability for endorsement.

MS. KAHN: And the overall suitability for endorsement. Does the measure meet NQF criteria? You can go ahead and start. You have 18 yes and 1 no.

CHAIR BROTMAN: Congratulations, we got through another one.

CHAIR SEPTIMUS: You know what's missing on that voting? We need the background music. You can put background music, can't you? We'll all go to sleep after lunch.

All right, 2080, "Gap in Medical Visit." This is also HRSA. And Michael.

MEMBER FARBER: This is again a similar measure to what we've been talking about. And the measurement is a little bit different in that it's looking for the last 6 months of the measurement year how many people still have made a visit in that last 6 months. And over how many people -- it's actually who didn't make a visit over the people who did. So that it doesn't have the same issues as the other measure in which -- of the newly enrolled.

But the issue is again that there are people with HIV that are lost to follow-up after being seen. So those who would be seen in the last 6 months of the year would have a greater issue of continuing and embarking

on the type of measures that they would need to get of CD4 counts and counseling. So that's the nature of the measure.

CHAIR BROTMAN: Any specific other issues related to impact?

MEMBER FARBER: Well I think that, you know, many of the studies that have been cited are all the same ones. There are 14 studies that have been cited in this on a meta analysis which basically -- the answer to them because they don't measure exactly what's in here. But what it is is that retention in visits leads to better outcomes for patients and as far as survival and also transmission because they have -- if they get on antiretroviral medication they then have a lower transmission rate. So those would be the reasons.

And again, this is a similar idea and that is where do you start measuring people for retention of visits. And this one is looking at where there's been a gap and they've

come back in a sense in the last 6 months of the year.

CHAIR BROTMAN: Peter?

MEMBER HAVENS: Just to again confirm with the developer that the intended population of study here would be providers who predominantly serve people with HIV. While that is potentially difficult to exactly specify you know it when you see it because you are funding it.

MS. MATOSKY: Yes.

CHAIR BROTMAN: All right. Aaron, I'm sorry.

MEMBER MILSTONE: Does that also apply for the medical visit? Because if you're looking at the facility it's not one visit with an HIV provider and then 6 months later with your obstetrician. Those are going to be -- do you have a way of identifying or specifying who the medical visit is with? Because there wouldn't be any data to support seeing your OB one 6-month period and your HIV doc the next

6-month period.

MS. MATOSKY: Our intent is that it would be used within a clinic and most often the obstetrician is not part of the clinic. It's usually an HIV clinic where it's just physicians, NPs, PAs.

CHAIR BROTMAN: Mohamad. I'm sorry.

MEMBER FAKIH: I was sure that the association is causal. I mean, all of these may be factors. You know, when we talk about visiting for 6 months, you know, it could mean they're within 6 months. Does it really mean that the presence in that office was related to improvement in health or you know, better HIV control or better outcome versus other factors?

CHAIR BROTMAN: Can you speak to that, measure developer?

DR. CHEEVER: So I think in this measure what we're looking at is people that did not have any medical care in that facility

for the last 6 months of the year. So, it's really the absence of care or evidence of that kind of specialty care is what we're trying to look at here.

MEMBER FAKIH: Does the absence of care in that facility for the last 6 months mean that that facility was responsible for worse outcomes? Is it causal? Do we have data about that?

DR. CHEEVER: We know that people that -- the studies that we cite are people that are not getting -- that aren't -- generally this is HIV care that we're looking at, that aren't getting HIV care do worse than people that are getting HIV care. In terms of causal as in -- I'm not exactly sure how to answer that or exactly what -- how to have causal inference in this.

MEMBER FAKIH: So the reason behind my question, you know, I think we have gone through so many measures right now and at one point we're going to ask ourselves the question

if we let's say have from 50 percent compliance of visits within the last 6 months to 90 percent do we think this is going to be impacting their care. I'm not saying that it wouldn't, but would -- you know, we are assuming because those that are there having the care are getting better outcomes. But it doesn't mean that population that is not having the care, that if they go to that office they will, you know, their outcomes will be any better. I don't know if I'm explaining it. Maybe they won't take their meds. Maybe, you know, maybe there are other issues. They don't have a house. They can't reach the pharmacy.

CHAIR BROTMAN: Tom, Doug and then Ed.

MEMBER GIORDANO: There is no way to randomize people to either stay in care or be out of care. So the causality is extremely difficult to prove. However, there are very consistent observational data showing that -- and pretty well-designed studies from very

large and ranging from small to multi-center large studies showing that if you don't have, if you're not retained in care that you are less likely to be prescribed ART, you're less likely to adhere to ART, you're less likely to achieve viral suppression and your survival time is shorter. So there's no -- is that causal? I don't know. But clearly if you're not in care you can't receive interventions to try to improve adherence to ART. You're not going to be prescribed ART. And so you're going to do worse.

Now, if you bring people back are they more likely to be in care, to get those things as a result? And in fact there's observational data from a SPNS project to suggest that yes, you can if you bring people back in care or if you keep them in care through interventions that they will, they can do better.

CHAIR BROTMAN: Thanks, Tom. Doug?

MEMBER CAMPOS-OUTCALT: This is the classic problem with observational data. Is it correlational or is it causational? And there are ways that you can assess observational data to have more confidence in it and one of which is to control for confounding variables and compare patients and so forth. I hadn't heard any description of that regarding the evidence that we've been presented. So, did the evidence report that was done, the meta analysis do that kind of assessment and if so how did they rate the final evidence?

CHAIR BROTMAN: Let's stay on impact right now. Ed?

CHAIR SEPTIMUS: I just had a question and maybe I missed this but I'm assuming that because the patient had a visit in the first 6 months that that indicates that the patient is in fact continuing to be followed by the same physician and therefore if he doesn't follow up in the second 6 months that's

a gap? I'm asking -- maybe this is dumb. I'm asking a question. How do we know that in fact that patient has decided to continue to be followed by that physician?

MEMBER MILSTONE: We don't but we seem to have ignored that in the last measure as well because we didn't talk about people that drop out of care, get incarcerated. So I'm hoping that's a small percentage in which case that's why you're saying it's okay because you're not expecting 100 percent.

CHAIR BROTMAN: Measure developer?

DR. CHEEVER: Yes, I think in fact we did acknowledge that even in this discussion, that people -- we need to have exclusions like people that are incarcerated, et cetera, and we took that out because it was almost impossible to code for it. I think at a jurisdictional level we've done very different kind of work where like across a state you have a better sense of in New York City if they hop from one provider to another.

Other places it's just less relevant because there's no place else to go.

CHAIR BROTMAN: We have to move on soon, but Kathleen quickly and Doug.

MEMBER BRADY: Just to follow up on that in terms of the extent to which that occurs. From data from the Philadelphia EMA I can tell you that less than 3 percent of people with HIV and AIDS get seen by multiple providers in a 12-month period. So overall it's small.

CHAIR BROTMAN: Okay. Let's go for a vote on high impact now.

MS. KAHN: Voting on 1(a) high impact. Go ahead and start. So we have seven high, seven moderate, two low and three insufficient evidence.

CHAIR BROTMAN: Okay. So that passes. Let's talk about the evidence.

MEMBER FARBER: Well, I think that the evidence is similar to the other studies in that all of them are looking at the continuation of visits, and that the

continuation of visits again were 14 studies that were subjected to a meta analysis and that in these studies that the continuation of visits resulted in many parameters of improved survival. That is, that resulted in improved survival. And that is getting more frequent CD4 counts. And that is many of them defined it as that was the issue of retention is whether you had CD4 counts within 3 months. So I think that the -- our group felt that the evidence was mostly high and there was one moderate.

CHAIR BROTMAN: Okay. Any comments on that? I think we've talked about some of this before. So let's go to a vote on the evidence.

MS. KAHN: Voting on 18 evidence. You can go ahead and start. We're waiting on one person. You have 13 for yes, the body of evidence meets the guidance, 1 for no, the evidence does not meet the guidance, and 3 for insufficient information.

CHAIR BROTMAN: Okay. So that

passes. Let's just briefly go to performance gap. Michael?

MEMBER FARBER: Well, we felt that there was certainly a lot of room for submitting this for quality improvement considering that the amount of retention was about 70 percent, 60-70 percent in most of the studies. Disparities were also noted, in females and minorities especially.

CHAIR BROTMAN: Any other comments? Okay, let's vote on performance gap.

MS. KAHN: Voting on 1(b) performance gap. You can go ahead and start. You have 6 high, 12 moderate, zero low and zero insufficient.

CHAIR BROTMAN: Okay. So that passes. How about reliability, Michael?

MEMBER FARBER: The group felt that the evidence was fairly reliable because of the equivalency of most of the studies showing the same direction of retention of visits

leading to better outcomes.

CHAIR BROTMAN: Any comments about the reliability? I think we have the results up there on screen for those of you in the room. Any specific comments? No? Let's vote on reliability then.

MS. KAHN: Voting on 2(a) reliability. Go ahead and start. We have 4 high, 14 moderate, zero low and zero insufficient.

CHAIR BROTMAN: Okay. And validity.

MEMBER FARBER: We didn't find any -- we felt the validity was generally high in this and that's how the group saw it.

CHAIR BROTMAN: Yes, go ahead, Peter.

MEMBER HAVENS: Again the question is raised about the focus of measure on HIV providers. In the prior discussion there was some statement that it could have come to the health system for another visit but you are

specifically talking about visits to a person in a clinic that routinely takes care of people with HIV.

MS. MATOSKY: Yes.

MEMBER FAKIH: Can you tell us how you reached the observation that it's highly valid?

MS. MATOSKY: So, as indicated in the measure submission form we used -- we reached validity through face validity. So we had a technical work group that designed this measure and went through a series of voting, rounds of voting for this measure. And it was found to be usable and feasible and have an impact on quality improvement.

And from there what we did was, because our technical work group consisted of 20 to 25 folks. What we then did was we had a series of webinars where we invited the Ryan White providers across the country to review the measure. We reviewed the measure during the webinar and sought input and feedback on

the measure.

And through our process we've had, I think we're now in data collection number 5 since last October. And we've had well over 130 providers from across the country utilizing this measure. And all of them had said that they found this measure to be easily implemented, easy to collect this data, easy to interpret and important to their quality improvement programs.

CHAIR BROTMAN: Any comments regarding that? All right, let's go vote on validity.

MS. KAHN: Voting on 2(b) validity. You can go ahead and start. We have 2 high, 14 moderate, zero low and 2 insufficient evidence.

CHAIR BROTMAN: Great, so that passes. Let's talk about usability.

MEMBER FARBER: I think this is -- we found it to be very easy to perform and to measure. Easy for providers to assess because

it's just one visit in 6 months. Can be easily done with electronic health records and without -- so that -- and it's been used in many studies already so that the proof of its ease has already been demonstrated.

CHAIR BROTMAN: Any comments before we vote? Let's vote on usability.

MS. KAHN: Okay, voting on usability. You can go ahead and start. We have 8 high, 10 moderate, zero low and zero insufficient.

CHAIR BROTMAN: Okay. Feasibility, Michael.

MEMBER FARBER: That's kind of the flip side of usability. If it's already being used there's a lot of feasibility to continue to use it. And there would be no reason to think that there would be a problem in implementing it for providers.

CHAIR BROTMAN: All right. Any comments before we vote? Let's vote on feasibility.

MS. KAHN: Voting on feasibility. You can go ahead and start. I think one more. Seven high, ten moderate, zero low and zero insufficient.

CHAIR BROTMAN: And finally suitability for endorsement.

MS. KAHN: Does the measure meet NQF criteria for endorsement? Go ahead and start. You have 18 yes and zero no.

CHAIR SEPTIMUS: Okay. Moving right along. The last in these suite of measures. Oh, look at the HRSA people, they're just so happy to leave.

(Laughter)

CHAIR SEPTIMUS: We'll get you back later. I guess -- no, they're finished it looks like. Okay, well thank you very much for your time. But we're going to have the Jenna and Bob show here as NCQA comes back. And Peter this is going to be yours I believe, correct? Okay. So would either of you like to make a brief comment about "HIV/AIDS Medical

Visit" 0403.

MS. WILLIAMS-BADER: Yes. So this measure -- I'd like to open this measure is included in stage 2 of the meaningful use program and it has also been adopted by the initial core set of healthcare quality measures from Medicaid-eligible adults. And it does align exactly with the National HIV/AIDS Strategy which defines continuous care as at least two visits at least 3 months apart.

You will notice that we have two numerators for this measure, one that's 90 days apart -- and two measures at least 90 days apart and the other two visits at least 180 days apart. That was due to some discussion among our experts about capturing patients that are not coming in for acute care, that you are seeing but that they wouldn't necessarily define as retained in care. The retained in care they think is best defined as two visits at least 180 days apart.

Also, the measure is not yet

included in PQRS so we weren't able to provide you with performance data from that program. And since the measure was just recently implemented in meaningful use and the Medicaid core set we don't have data for that either. But we did present data from the National HIV/AIDS Strategy. That's all I have.

CHAIR SEPTIMUS: Okay. Peter, let's talk about impact.

MEMBER HAVENS: Thank you very much. You'll notice that for the last few I asked the same question over and over again because this measure is really different than the intent of the prior two measures which we've just endorsed. This measure is for patients who are in HIV care and who within a 12-month monitoring period have had two medical visits with a minimum 90 or 100 days.

Data are presented to suggest the importance of getting patients into care and keeping them in care but compelling data are not presented in the summary to suggest that

the identified visit frequency or duration of follow-up of one year are optimal to make this assessment. So conceptually this has high impact but operationally it might be discussed if the 2-year time line as outlined in one of our prior reviews might be a more appropriate measure in that period for care of patients with a chronic illness.

Not to focus on it too much here because we may get to it more in validity but to again point out that this is for visits to any practitioner, not just a -- well, let me ask the developers. Is the intent here that the practitioner of record being counted for a follow-up visit be an HIV -- person who generally cares for people with HIV? It lists a pediatrician or an OB/GYN in the list of practitioners which would seem to be somewhat different than the HRSA measures we just reviewed although I understand the problems with identification of those people.

CHAIR SEPTIMUS: Bob, you want to

comment on that?

MR. REHM: Sure. I think that we recognize that HIV care especially if we think 5 years down the road is going to be provided probably at a different level than it currently is and they'll be more integrated into primary care writ large. And I think the intent of this measure is to capture patients whoever they see for primary care. And we would think that they would be able to provide the kind of care that we're talking about here in terms of having two office visits within the year.

So, from a definitional standpoint I'm not sure how because I'm not close up and personal with the HRSA measure how in fact -- it's one thing to have a measure intent and they're focused on their clinics, but if we think that we're trying to basically develop a nationally endorsed measure for broad utility I don't know how they define that in the measure.

So my sense is it's probably hard to define because where would you do that. I mean you could have an attribution logic, you know, but I don't think anyone really wants to go there with that because it's complicated.

So you're correct, we think that this is for primary care practitioners. And I come in with diabetes, I come in with HIV, I come in with CHF, a variety of different things, much of which can be managed without necessarily going to see a specialist or a specialist augments that service but the primary care really, that's the focus of a lot of our measures. So we're comfortable with that.

MEMBER HAVENS: Thank you for that clarification.

CHAIR SEPTIMUS: Anyone want to comment then on the impact of this measure? If not then we're ready to vote.

MS. KAHN: Voting on 1(a) high

impact.

CHAIR SEPTIMUS: Did I miss him? We don't have to vote. We can wait. Go.

MS. KAHN: Voting on 1(a) high impact. I'm not sure who is at their seat and who's not anymore. So we have six high, nine moderate, zero low and one insufficient.

CHAIR SEPTIMUS: Okay. Peter, the evidence, please.

MEMBER HAVENS: There are no randomized trials so the evidence can be at most of moderate quality. Many of the guidelines cited suggest expert opinion as the quality of their evidence but there are cohort and case control studies showing the benefit of visit frequency as a marker of adequacy of care.

CHAIR SEPTIMUS: Any other comments then about the evidence? Aaron.

MEMBER MILSTONE: I was just curious if there's any evidence that seeing an obstetrician twice a year improves outcomes

in patients with HIV.

CHAIR SEPTIMUS: Kathleen.

MEMBER BRADY: You know, it's been awhile since I looked at this data but it's not -- it's seeing a provider who is familiar with HIV care and having a certain volume of patients who have HIV that make you proficient in treating HIV.

MR. REHM: First on his question, I'm trying to remember it now. It was -- oh, OB/GYN. Quite often when we are looking at primary care and this is kind of on the NCQA side of doing measures for 21 years OB/GYNs often are the primary care provider of choice by many women. And so it's not -- it's not that they're going in for necessarily an OB/GYN visit, it's they're using their OB/GYN as a primary care provider. So that's trying to be inclusive rather than exclusive.

Kathleen, to your question, again the way you'd get around that is applying some attribution logic that says X percent of my

patients have HIV diagnosis. I don't think we've seen many HIV measures come forward that suggest that that's tenable. So I mean, we're kind of in a position between transition between HIV care being provided by as you characterize HIV I wouldn't call them specialists but people who are highly tuned into this practice as opposed to again we are sensing that practice, that primary care for these patients is going to be provided by a broader spectrum of clinicians.

CHAIR SEPTIMUS: Mary?

MEMBER BLANK: Do the two visits have to be with the same provider or same specialty?

MS. WILLIAMS-BADER: No, the measure does not require that. So, for example, the measure that would be used in meaningful use, the eligible professional that's reporting the measure would just need to have access to the information that the patient has had two visits in their EHR. So

if they do know that the patient has had two visits then they wouldn't necessarily have to be with the same physician.

Now, if it's in a system where the physician might not know if the patient has seen another provider then you would have to -- I guess it would have to be with the same provider.

CHAIR SEPTIMUS: Tiffany?

MEMBER BLANK: How does that get into continuity of care though if it's not a particular provider that's following them?

MS. WILLIAMS-BADER: I think that what we would be picturing for again the measure being used in meaningful use is that it's likely the other provider, if the information is available in their EHR is a provider in the same clinic or someone whose information they would have available in the EHR. So.

MR. REHM: Yes, I could use an example although EHR is not my zone as you know. I go to see my primary care physician. They

realize that my HIV diagnosis maybe among others requires me to go to their HIV specialist. And in that setting then you're actually capturing continuity because you're capturing the referral and the activity within that. So that would be an example where even though it's two different providers I would characterize that as continuity of care.

MEMBER BLANK: Would that referral take 6 months for the numerator, the second numerator?

MR. REHM: I was just using an example of where you could have two providers providing care and it would not be discontinuous.

MS. WILLIAMS-BADER: I guess it's unlikely that a provider would have information about visits with other providers unless it's an integrated system. So then the patient might be seeing several providers within that system but it's an integrated system and that's how the information is available in the EHR

in the first place. So if you are an HIV patient and your regular primary care doctor is not available when you come in for a follow-up visit if you see someone else in that setting then you would -- then an eligible professional reporting on the measure could get credit for that. And it would be continuity of care because it is two providers within that same setting.

CHAIR SEPTIMUS: Tiffany?

MEMBER OSBORN: So, my question would be we've already passed a couple of these that talk about making sure that there's continuity of care and number of visits per year and all of this. So I guess my question here would be is there data specifically relating to this 90 and 180 days that makes us need to consider this any differently than the measures that we've already passed? I mean, what is it about the 90 and the 180 days versus the two visits in a year versus the first 6 months and the last 6 months?

MS. WILLIAMS-BADER: So first of all this measure is an already-endorsed measure. So actually this measure has been around since 2009 and has been endorsed since then. The HRSA measures are new measures that are being presented today.

The second, we do think that there is a difference as far -- and while we don't have the data specifically for 90 days or 180 days, like I said, the 90 days does align exactly with the National HIV/AIDS Strategy. So if -- I think one of the things that we really try to do at NCQA when we're developing measures is try to align as much as possible with national programs so that there is some continuity across all of those programs as well and you don't have different numbers from different programs, or different goals that you're striving for.

And like I said, the 180 days was really to further delineate those patients that are really retained in care that are coming

back throughout the year.

CHAIR SEPTIMUS: Just a quick follow-up. So how do we know the visit was for HIV care?

MR. REHM: It's not required. It's not required that the visit be for HIV care.

CHAIR SEPTIMUS: So I guess I'm just asking, we'll have other people comment but I'm -- that would be problematic. But Aaron? I'm sorry, Kathleen.

MEMBER BRADY: That's all right. No and this goes back to an example. So I work at the University of Pennsylvania Health System. So if a patient comes to see me for HIV care but then, you know, as you mentioned before goes to see their OB/GYN who doesn't treat their HIV, they're just getting their annual pap smear I have an integrated health record so I can see that they went to the OB/GYN but that would meet your measure. If it was correct.

MS. WILLIAMS-BADER: Right. At this time in EHRs it's very difficult to ascertain the intent of the visit and to be able to capture that reliably across all EHRs. So I think that might be something to consider for a future state of the measure where you would want to make sure that it's for HIV care. So that's just -- it's something that we've definitely discussed and considered. It's too difficult to capture at this time.

CHAIR SEPTIMUS: Now, Aaron.

MEMBER MILSTONE: Thanks. No I think that was exactly -- so the previous measure was very clear in that it was targeted to where -- I think we all thought it was clear to say it was targeted toward HIV providers. Here it's targeted more broadly to primary care providers so if someone comes in with vomiting to see a primary care doctor and HIV 6 months later. So I still feel like that comes back to the evidence, is there evidence that those other visits for other HIV-unrelated

issues is going to benefit the patient.

CHAIR SEPTIMUS: Mary, did you have something? Go ahead.

MEMBER GIORDANO: On the issue of the 90 days and the 180 days I would agree with the developer that this is consistent with other standards of care around HIV. And that a -- there are data showing that people do have worse outcomes if they have fewer -- worse retention in care and that is measured in variable ways.

But clearly if you don't have at least 2 visits at least 90 days apart you're going to do worse. I think there are a number of ways to -- but with an HIV specialist, with an HIV provider. I think that caveat is important to note.

CHAIR SEPTIMUS: So, and so actually I think we've sort of hit on a key element because it doesn't specify that. And so I think that's, I think from the evidence standpoint that's going to be, for some of us

may be a barrier.

MR. REHM: Well, yes. In terms of a level playing field, and again I haven't looked at the HRSA measure but I have to -- I'm not sure that they're specifying in the measure specification what the definition of an HIV specialist is. To me it's as open a book as ours is. It -- the intent is one thing but again, the intent around measures that are used in HIV clinics and the like is one thing. This is -- this committee is voting for nationally endorsed measures to be used in a broad setting so I'm not sure I recalled seeing the definition of that practitioner. I understand the intent. The intent is different.

CHAIR SEPTIMUS: I think the question is whether the visits are for HIV care, not whether it's an HIV specialist. I think that's -- that's I think what some of us are voicing as concerns. But I don't want to take away other people's time to talk. So did you

have a quick answer, Jenna, on that?

MS. WILLIAMS-BADER: Yes. Again I think the intent would certainly be that it's for HIV care but I don't know that there is a way to specify right now that the visit is for HIV care. You can look for a diagnosis of HIV for that visit but that doesn't mean it's the primary reason for the visit.

CHAIR SEPTIMUS: Adam?

MEMBER THOMPSON: Yes. One example I just want to give to consider is especially in rural care where right now we're building the capacity of primary care providers to pull labs and interpret those labs, and then they're being seen in an infectious disease specialist once a year. And there's no guarantee that those two providers have an integrated system. Yet it would be two separate visits across two providers, both providing HIV care, though one would not be seeing a predominantly HIV population nor necessarily be an HIV specialist. So it's just

something to consider.

CHAIR SEPTIMUS: But the visit could still be for HIV care.

MEMBER THOMPSON: It depends on how you would define that. I mean, if pulling labs is considered an HIV visit rather than seeing a specialist who knows how to diagnose some complex opportunistic infections then yes. But if not you would need to have some higher level of capacity.

CHAIR SEPTIMUS: Michael and then Tiffany.

MEMBER FARBER: I wanted to say that the retention in many of the studies was defined really not by visits but by CD4 counts being performed.

I'd say also traditionally, you know, years ago there were a lot of non-infectious disease doctors who saw HIV patients. But with the explosion of antiretroviral therapy and the complexity of it there are even many infectious disease

doctors who don't feel that they're specialized anymore in HIV. So I guess my comment about the issue of what the visit is for, you know, especially in the medical home there would always be an attempt to try to get a network of that HIV provider.

But it isn't known at all that the visit would be for HIV at all. It could be just for bronchitis and the person might not at all address the issue of labs. But of course that would be optimal if they did and then just like any other generalist that they make a referral when they realize that the complexity of the problem is beyond their expertise just like referring to a cardiologist when there's coronary artery syndrome.

CHAIR SEPTIMUS: Tiffany?

MEMBER OSBORN: I just want to make sure that I'm clear. I mean, because it was brought up that this is a measure that's already been endorsed, is coming back. And we've discussed that regarding the specific time

frames 90 and 180 days there's not a lot of evidence to support that or the fact of seeing a non-HIV provider versus an HIV provider for the two subsequent visits. Do we treat this relating to evidence any differently because it's a measure that's coming back or was previously endorsed?

MS. WINKLER: No. The criteria apply equally to all measures new or previously endorsed.

MEMBER OSBORN: One thing to clarify though because I want to make sure it's clear. When I go back and look at the HRSA measures they're specified very similarly in that it calls for a medical visit, calls for patients with a diagnosis of HIV/AIDS. The assumption I think with all of these measures in front of you is that the ones, the clinicians who would be measured on these would be the ones who typically treat these patients but there's no way right now that I am aware of to be able to determine that visit, that

clinician provider in that way is solely treating HIV/AIDS patients and that the visit itself is for that diagnosis.

And so it's a general problem, I think a challenge across all of these measures, not just this one measure in particular. I want to make sure that's understood by everyone.

CHAIR SEPTIMUS: Okay. Any more discussion then about the level of evidence? Okay. If not we will vote.

MS. KAHN: Voting on 18 evidence. You can go ahead and start. Everyone press it one more time. You have eight for yes, the body of evidence meets the guidance, four for no, the evidence does not meet the guidance, and four for no, insufficient.

CHAIR SEPTIMUS: It's a tie.

MS. KAHN: We're missing two votes also.

CHAIR SEPTIMUS: So Ray's not on the call?

MS. KAHN: We should have 18 right now.

MS. WINKLER: There should be 17 is my count. Then let's try and do it again.

CHAIR SEPTIMUS: Okay, let's re-vote then.

MS. KAHN: Okay, you can go ahead and start.

CHAIR SEPTIMUS: Go.

MS. KAHN: Can we do it again?

CHAIR SEPTIMUS: Well, but this is -- but now we have a majority on the evidence. So somebody changed their vote. Sixteen. But now it's not a tie so somebody changed their vote. I say we need to just move on and we'll go with the opportunity gap.

MS. KAHN: Just for the record it's nine yes, the body of evidence meets the guidance, three for no, the evidence does not meet the guidance and four for insufficient information.

MEMBER HAVENS: Concerning the

opportunity gap, Section 2b.5 states that 73 percent of patients have at least two visits per year at least 60 days apart, identifying that there would be opportunity for improvement.

CHAIR BROTMAN: Any discussion on that?

CHAIR SEPTIMUS: Well, we're just delighted to vote. So let's vote.

MS. KAHN: Voting on 1(b) performance gap. You can go ahead and start. I think someone's battery died. Zero high, 13 moderate, 1 low and 2 insufficient.

CHAIR SEPTIMUS: Okay. Now we're going to talk about our two favorite indicators, reliability and validity. So, starting off with reliability.

MEMBER HAVENS: In terms of reliability, again we note that HIV specialty care is not required of the visit type but in the last number of measures that we have looked at if this were applied to HIV specialty care

providers then that is the visit type that would be counted.

And looking at EHR versus manual calculation of performance at 91 percent versus 95 percent were identified as meeting the goals. So this is within 4 percent of each data type suggesting reproducibility of manual versus EHR calculation.

While we're talking about the combination of reliability and validity, face validity was assessed by six experts who agreed 100 percent that this was a good measure of quality of care.

CHAIR SEPTIMUS: Boy, that's pretty unusual.

MR. REHM: Actually it was 4.67 on a 5 scale. One hundred percent though voted, so.

(Laughter)

MS. WILLIAMS-BADER: One hundred percent strongly agreed or agreed that the measure is a good quality care measure.

MEMBER HAVENS: My god. You know, I tried to present their data as positively as I could.

(Laughter)

MEMBER HAVENS: I got creamed for it.

CHAIR SEPTIMUS: Any other discussion about reliability? Seeing none, we'll vote.

MS. KAHN: Voting on 2(a) reliability. You can go ahead and start. You have 1 high, 11 moderate, 1 low and 3 insufficient.

CHAIR SEPTIMUS: Okay. So let's go onto validity.

MS. KAHN: Voting on 2(b) validity. You can go ahead and start. We have zero high, nine moderate, three low and four insufficient evidence.

CHAIR SEPTIMUS: Did you want to comment on something before we go to usability, Peter? Please.

MEMBER HAVENS: So, while this passes on that criterion I wanted to point out to the developers of these and the other measures that they should not expect that if they cannot begin to identify what provider they think is important in the outcome of care that they should not expect endorsement of these measures when they come back to a body such as this in the future.

If you're going to apply this to everybody in the country it is your responsibility to show data that it measures something that matters. And if this comes back in 3 years without better data if I'm on this committee I will be glad to comment more specifically on issues of reliability and validity in measurement of this outcome. Thank you.

CHAIR SEPTIMUS: I think we've heard a lot about better how to define a visit and what the purpose of it is. Aaron?

MEMBER MILSTONE: Can I comment on

usability now?

CHAIR SEPTIMUS: Well, Peter has another chance to comment on usability. If you'd like. Okay, he yields to you, Aaron.

MEMBER MILSTONE: So yes, I feel similarly. I have trouble with how this is being applied currently in its face validity, in its usability in terms of understandable and useful for public reporting.

So again, I think a person who goes to see their primary care physician and is managed for HIV and then goes to see their obstetrician 6 months later for a pap smear, that's not the intent of why we're trying to retain people in care for HIV. So to me that is not meaningful and useful.

And I think that it's fine to say that we want to see this improved in 3 years but we're endorsing this now for the next 3 years which means it will impact -- there will be implications of this. And I think people need to take that seriously in their

considerations.

CHAIR SEPTIMUS: And just to remind the committee, usability is not a stop vote. So we don't have any other stop votes until we get to whether or not the measure is suitable for endorsement. So just to let everybody know. So Peter and then Tom.

MEMBER HAVENS: And I appreciate your comments but I'm not sure I agree with them. And I think this complexity of identifying who you should really see is really complicated. So I'm not saying it shouldn't be the pediatrician who -- or the family practice guy in the rural area. This may be completely reasonable. But we need to be studying this. You know, when I see a patient for HIV care if he goes and gets vaccinations from a pediatrician that is a really important part of routine care and it's probably cheapest done at another site. So for me to say that I want some clarity on the measurement is not because I don't -- that I don't agree

with (a) the problems that have been identified, and (b) that maybe this kind of retention is important but we need to be looking at that over time since we're spending a lot of money now to make it a part of meaningful use.

CHAIR SEPTIMUS: Tom?

MEMBER GIORDANO: Is there -- this is I guess more a question about these measures in general. Is -- does the developer matter? So, if HRSA develops --

CHAIR SEPTIMUS: You're feeling really beat up now?

(Laughter)

MEMBER GIORDANO: No, I --

MR. REHM: We're out of here.

MEMBER GIORDANO: Once these things are sort of -- are blessed or whatever the proper phrase is for this endorsement can anyone pick them up and use them or is it still sort of the developer's cadre of clinics that ends up using them?

You know, if HRSA has a measure that's endorsed and wants to push the people it pays to provide HIV care to use that measure I think that makes sense. If the -- on the other hand, if the NCQA has a measure endorsed does it mean that it would be potentially used by everyone, or can they also say well, we just want our HIV providers to use it? I don't understand that.

MS. WINKLER: Okay. That was sort of the context I was trying to explain to you at the very beginning of our meeting yesterday was the intent of NQF endorsement is to identify measures that can be used quite broadly on a national basis. They are openly available for any potential end user and we are -- encourage and are looking for the measures that are going to be most widely adopted.

So, on the one hand when you say does the developer matter it matters very much because they maintain the measures so that they stay current. But in terms of ultimate end

users not necessarily. And that's why the measures and one of NQF's roles is to evaluate the measures, endorse the measures. We have our database that's available on our website for people looking for measures to use. They can come, it's a resource, they can find the specs, they can get all the information they need to potentially put it in whatever program they're putting it in.

And as you saw, I showed you the pie chart of the uses of the various measures that NQF has endorsed. You can see that they're used in a wide variety of different kind of public and private programs.

CHAIR SEPTIMUS: Okay. Yes, Mary.

MEMBER BLANK: From a health plan perspective we endorse measures, we pull them into our models and we go back to the developer if there's any questions on the specifications or how they work through something. Regardless of the developer if it's something that we want to put focus on in one of our

pay-for-value programs we'll pull it in.

CHAIR SEPTIMUS: Bob has a comment and then.

MS. BOSSLEY: Okay. I mean, to me I think you should assume that the measure -- any measure that is endorsed could be used by anyone. So it may be used -- the HRSA measures may actually be used first by HRSA, but it's very likely that, and I think it would be their goal as for all developers anyone else will uptake it. It's the same for an NCQA measure, a CDC measure, any of the measures we see here, it's really anyone can use that measure. That's the goal if they want to.

CHAIR SEPTIMUS: Okay.

MR. REHM: And just to add kind of a reality check to this because I think there's a big fear of unintended consequences, people being measured that really shouldn't be.

In truth when these measures go out, whether they're NCQA without NQF endorsement because remember there used to be a world before

that endorsement was a dominant feature of our life, people would adopt a health plan measure, modify it for clinician groups, do it in a regional collaborative, don't call it HEDIS because they'd be violating the specifications but they would use those and there would be utility in that. And they would use those for targeted areas. I don't think anyone's going to be hunting around and saying gee, let's use this measure and focus on the OB/GYN community because it's available and they happen to be listed as a provider who could provide that service.

So, I think that we don't want to overreach the fact that the NQF measures if they go out there, people look at them. If they have particular value to their operations or maybe to their pay-for-performance you know they may focus, but they're not going to focus it on things where they're going to get clinician, you know, pushback and anger. You know, it's -- there's too much stuff going on

right now. And so I think there is a logic to what gets used and what doesn't.

And both Heidi and Reva are absolutely correct, national endorsement means they're out in the open. They're in the portfolio. You can use them or not. It's paint set. Do you want to have more colors or do you want to have less? I hear less is better you know.

CHAIR SEPTIMUS: Okay, Mary, you have another comment? Okay. Tiffany?

MEMBER OSBORN: I think all of this is probably getting off the actual point of what we're supposed to be doing but if this ends up coming for -- ends up going to CMS then we can't pick and choose. This is applied to everybody, right?

MS. WILLIAMS-BADER: Well, it is in CMS's program. It's in stage 2, it's in the final rule. So this is going to be used for stage 2. CMS can actually make some modifications and have made modifications to

NQF-endorsed measures when they use them in programs like PQRS. They actually do.

So, again, it's for the meaningful use program these are -- the providers that are participating get -- it's an incentive program first of all, it's voluntary, and then they get to select measures that they want to report. So it would not make much sense for someone who doesn't provide regular HIV care to report these measures because honestly their rates will probably be low. These measures are likely to be picked up by providers that are providing HIV care.

CHAIR SEPTIMUS: Okay. Any other comments about usability? So not seeing any, let's vote.

MS. KAHN: Okay, voting on usability. You can go ahead and start. We have one high, six moderate, seven low and two insufficient information.

CHAIR SEPTIMUS: All right, well this one then, with my arithmetic from grade

school this one would fail usability. Nine versus -- I didn't say it was a must-pass, I said it failed usability.

All right. Feasibility. All right, any other comments on that or should we vote on feasibility? Okay, let's vote on feasibility.

MS. KAHN: All right, voting on feasibility. You can go ahead and start. Can we have everyone press it one more time? Zero high, eight moderate, six low, two insufficient information.

CHAIR SEPTIMUS: It's a tie. All right, now, the last one. Is this suitable for NQF endorsement? This is a simple yes or no.

MS. KAHN: Does the measure meet NQF criteria for endorsement? You can go ahead and start. We have 6 yes and 10 no.

CHAIR SEPTIMUS: Okay. I think there's some take-home messages I think for our developers on this one, so. All right.

Well, the next one -- we're just going to keep going.

CHAIR BROTMAN: This one is Ed.

CHAIR SEPTIMUS: It is me which is 0408 "HIV/AIDS TB Screening." This is also NCQA.

MS. WILLIAMS-BADER: I don't have any comments to make about the TB screening but I did want to clarify about the chlamydia, gonorrhea and syphilis screening measure that this used to be -- when the measures were originally endorsed it was two measures. And recently we thought it made sense to combine them and provide a better picture of the STD screenings that patients with HIV are getting.

I also wanted to point out that there is a fairly large gap between the automated and manual performance rates for the chlamydia and gonorrhea testing measure. The reason why there is that gap is because there was a technical glitch in the EHR where this measure was being tested that was loading lab

data into an incorrect field. So, when you actually look at the syphilis screening measure which would also rely on laboratory data you'll see that there is a much smaller difference between the automated and manual performance rates. And I think you can take that into consideration in that if the chlamydia and gonorrhea data was being loaded into the correct field that there would actually be a lot more agreement between the manual and automated performance rates. Thanks.

CHAIR SEPTIMUS: Thank you, Jenna. That was noted on the workshop call. Okay, this is mine. So we're going to talk about impact. I think we can go through this fairly quickly.

I think most of you know that HIV and TB don't go well together and that people with latent disease have a much higher risk of going onto develop active tuberculosis and all the secondary public health issues surrounding that. So I'll stop there because

I think the work group agreed that this certainly has a high impact in terms of care and public health.

CHAIR BROTMAN: Any comments?

MS. KAHN: Voting on 1(a) high impact. You can go ahead and start. You have 11 high, 4 moderate, zero low and zero insufficient.

CHAIR BROTMAN: So that passes. Let's look at the evidence.

CHAIR SEPTIMUS: Lots of things are provided in here in terms of evidence. There is one randomized controlled trial. There are a number of practice guidelines that are appropriately graded.

I think when we start talking about, a little bit later on we'll talk about -- I'm not sure -- well, I think it's probably appropriately tested here. As most of you know with low CD4 counts obviously the reliability of the tuberculin skin test is not very reliable. There are interferon

gamma-releasing assays which probably are a little bit better. This document calls for either one. And of course the real challenge in this is that there has to be some clinical judgment in patients who are at high risk who are exposed. Independent testing is recommended to receive prophylaxis.

So, one of the challenges I think in terms of the evidence is that yes, it's a good idea to do this in terms of the impact but in terms of the -- in terms of the testing itself the testing has significant limitation, applying it to this population. So I think I'll stop there and see if anybody else wants to comment on the evidence.

CHAIR BROTMAN: Tom?

MEMBER GIORDANO: To reply to Ed's comments I agree with him that the testing isn't perfect but it's what we've got. And it still is a significant public health problem, especially in persons born outside the U.S. So it may not be ideal but it is the best we

have.

CHAIR BROTMAN: Any other discussion on the evidence? Okay, let's go for a vote on the evidence at this point.

MS. KAHN: Voting on 18 evidence. You can go ahead and start. So you should have one more person. So we have 13 for yes, the body of evidence meets the guidance, 1 for no, the evidence does not meet the guidance, and 1 for no, insufficient information.

CHAIR BROTMAN: Okay. So that passes. Let's just briefly talk about the performance gap.

CHAIR SEPTIMUS: Based on the document and also the literature there certainly is -- there are a lot of people who do not get these tests done so I do believe there is a significant performance gap.

I don't think there was anything about -- let me just double-check about disparity. I'm sorry, I should remember that. But I think the same thing we mentioned about

disparities about the other ones apply to this. Yes. I can't read that. What page is that?

CHAIR BROTMAN: It's up on the screen, 1b.4.

CHAIR SEPTIMUS: Only she can read this. Not stratified by patient groups or cohort. I'm sorry, I actually had it starred and I forgot it.

CHAIR BROTMAN: Okay. Any discussion on performance gap?

CHAIR SEPTIMUS: And the rate is low. There clearly is a gap in care for this measure being only 68 percent, so.

CHAIR BROTMAN: Okay. I think we're ready for a vote.

MS. KAHN: Voting on 1(b) performance gap. Go ahead and start. Eight high, seven moderate, zero low and zero insufficient.

CHAIR BROTMAN: So that passes. Now we're onto Ed's favorite portion, reliability.

CHAIR SEPTIMUS: I'm going to sort of probably take both of these together because they sort of overlap. The data sample, they use automatic electronic health. And they also did manual calculation in performance as well.

This is where the testing, there was a significant difference of 20 percent between the automated and the manual. The other thing is it's very difficult to capture because there's lack of standardized fields between the result interpretation, is it positive or been treated, or whether or not it's been asked for. It's only available primarily in the paper medical record.

So I think this is one which is sort of -- we'll get to I guess the feasibility later but it's going to be labor-intensive. There is a gap between manual and automated. And it's hard to capture this information. It's very inconsistently captured. So I -- and the work group also discussed this but in terms

of reliability and validity it's a problem with capturing the information.

CHAIR BROTMAN: Any comments for reliability, validity? If not -- oh, Peter, go ahead.

MEMBER HAVENS: What are we supposed to do with this information? I mean, does this -- if this is supposed to be used in an electronic health record what difference in data capture is reasonable from that perspective? If the EHR misses 20 percent in terms of a performance measure that's actually 40 percent overall, you know, out of the 50 percent who make it. So it's a big percentage of the overall issue.

There's a couple of ways around that. One is to not allow a PPD which can't be captured in the EHR but only allow IGRA testing or -- I'm just interested in how people would approach this or if this is okay.

CHAIR SEPTIMUS: I personally think the interferon gamma assay for this

population may be better. It is more expensive which is another consideration. Almost every practice guideline, and correct me if I'm wrong, has either/or as mentioned in the guideline. So, although interferon gamma-releasing assay has many attractive features, it's probably more easily captured in the electronic medical record, it doesn't require someone coming back to have it read by a trained individual, but it is more costly and right now guidelines say for either/or.

MEMBER BEAL: This is Jeff. I want to mirror that's the truth and in Florida the standard of care has become to try to place a PPD in a Ryan White population but if they don't return, go to the IGRA. And that would be missed by this current.

MEMBER HAVENS: No, that would be captured if they came back. Then you'd capture the IGRA when they came back for that. So that would still be okay.

MEMBER BEAL: All I see is positive

PPD. I don't see IGRA. Did I miss that? I understand that it's a definition of a TB screening test but I don't think that's specifically noted in the inclusion of the -- am I missing it? Tuberculin skin test in the numerator.

MR. REHM: It is included. We'll find out where it was specified as such. Because it was definitely discussed.

CHAIR SEPTIMUS: Documented TB screening was performed and results interpreted, at least one since the diagnosis of HIV.

MEMBER BEAL: I'd check that because the numerator says tuberculin TB screening test. I just think that that would be interpreted as a TB skin test but I appreciate it if it's not. Thanks.

MEMBER HAVENS: There's a note on page 9, a technical note that identifies that an IGRA is adequate for screening.

MEMBER BEAL: Excellent, thank

you.

CHAIR SEPTIMUS: Okay, ready to go on reliability and then we'll do usability right after that? Looks like we're ready.

CHAIR BROTMAN: Let's vote.

MS. KAHN: Okay, voting on 2(a) reliability. Go ahead and start. We have two high, six moderate, five low and two insufficient.

CHAIR BROTMAN: That passes. Onto the next. We're doing validity.

MS. KAHN: Voting on 2(b) validity. You can go ahead and start. One high, seven moderate and seven low, zero insufficient.

CHAIR BROTMAN: So that passes. Okay, onto the next section. Let's go to usability.

CHAIR SEPTIMUS: Under usability the measure is not currently used for public reporting. However, NCQA will submit NQF-endorsed measures for PQRS for consideration. And the TB screening is used

by HIVQUAL indicating the measure of this will focus on meaningful and useful for public reporting.

CHAIR BROTMAN: Any discussion? Kathleen.

MEMBER BRADY: I mean, my major concern is something we've talked about with other measures and that's the fact that since this is once since diagnosis there may be a lot of historical data that does not end up in an EHR and therefore gets missed.

CHAIR BROTMAN: Good point. Anybody want to comment on that or another comment? All right, let's vote on usability then.

MS. KAHN: Voting on usability. You can go ahead and start. We have zero high, 10 moderate, 4 low and 1 insufficient.

CHAIR BROTMAN: Okay. Onto feasibility.

CHAIR SEPTIMUS: Not much in terms of feasibility. They're not aware of any

unintended consequences related to this measure. So I think the same applies to this as usability. There does not appear to be any unintended consequences by what the developer has reported.

CHAIR BROTMAN: Aaron, do you want to make a comment?

MEMBER MILSTONE: Sure. I'm still unclear as to how the data on interpretation is going to be captured broadly, how that would impact the feasibility.

CHAIR BROTMAN: Any comments from the measure developer?

MR. REHM: So like you mean terms like positive PPD reported? You know, I think this is the classic where we are linked to the vendors and their capacity to -- and they are certainly improving recently to track the quality measures that are out there and begin to think about how they can establish those fields. And I think that's that, you know, that's where we're at.

If we believe that EHRs tend to move in groups and that there isn't one that gets really, really good at one little thing I would imagine that they would move together in a way. So in terms of comparability even though it's not capturing everything it's capturing what it can capture at about the same degree.

I know that's not much comfort but I think that's -- we can't really influence from a developer standpoint. I think we try because we meet with the EHR vendors all the time and as does ONC and say look guys, we have these meaningful use measures, you know, and can you please adapt your systems to better reflect what we're trying to capture.

MEMBER MILSTONE: That's terrific. Usually those meaningful use measures follow or are based on data that's been shown to be valid. But we're saying that we don't have the validity yet, right, to where if we have that in the EMR, if that's developed then we can show that it's a valid measure. But I feel

like we're putting the cart before the horse by we're creating a measure to drive vendors to incorporate that field into the medical record so it can be captured. But right now I'm concerned that with what people have it's going to be hard for people to capture whether that's been done or not.

MR. REHM: Yes, we appreciate the point. I think we're creating the measure because we think TB testing is important for the population, you know, and that given the ascendancy of EHRs and that this was tested in that setting to a moderate degree of success that's where we're at and understand the gap and recognize that. I don't know how we close it without that cooperation. We didn't create the measure to get EHR vendors to do better, we created the measure because it's an important public health arena and an important area to measure.

As we develop better capacity over time and the CHR landscape which a lot of people

thought was going to be a panacea and solve all problems and we know that that's not the truth. Should that keep us back from specifying it and putting it out there? And I think if we go back to measure development 15-20 years ago we understand that some of the measures in retrospect look pretty simple and kind of boring and not terribly helpful but at least we build on those. So I think that's the spirit within which we're putting this measure forward.

CHAIR BROTMAN: All right, any other discussion? Let's vote on feasibility.

MS. KAHN: Okay, voting on feasibility. You can go ahead and start. You have zero high, six moderate, six low and three insufficient information.

CHAIR BROTMAN: And ultimately let's vote on suitability for endorsement.

CHAIR SEPTIMUS: Well, although this is not a stop measure it's a negative response.

MS. KAHN: Does the measure meet NQF criteria for endorsement? You can go ahead and start. I think we're missing -- there we go. We have nine yes and six no.

CHAIR BROTMAN: Okay. It passes.

CHAIR SEPTIMUS: We have one more measure and then we're going to try to wrap things up and get everybody on their way. This is 0409 "HIV/AIDS Sexually Transmitted Diseases." I think our developer has already commented on this. And I know Kalpana is going to discuss this.

MEMBER RAMIAH: Sure. Last but not least measure. This is very similar to the TB screening document here. And the numerator is patients who have received screening for all three STDs, chlamydia, gonorrhea and syphilis, at least once since the diagnosis of HIV.

So the two points here is one, screening, the word screening was discussed in the subgroup as should it be screening or

should it be serological testing more clearly. And the second point here is about at least once since the diagnosis of HIV whereas the recommendation that the coding is actual screening. So that was a disconnect.

Do you want to comment on that now before we move on?

MS. WILLIAMS-BADER: Sure, happy to. As far as the screening, yes, for when -- the e-specification of this would clarify that it would be actual tests for the -- for chlamydia, gonorrhea and syphilis. And in line with probably other measures that we've created that results need to be present as well. That's generally the bar NCQA has set for our e-specifications and lab tests in e-specifications.

As far as the annual is considered we did have a lot of discussion about this with our own expert panel. I think there were experts who did believe that annually was appropriate if the patients are sexually active

but that it might not be appropriate for all patients, particularly those that are not sexually active and we certainly had some experts who said that not all of their HIV patients are sexually active.

And identifying sexually active patients is very hard to do consistently, reliably or validly right now. So that was why that criterion was not added. I think we would be open to annual if the group here feels strongly that it should be annual instead of once since diagnosis.

CHAIR BROTMAN: Any specific discussion? All right, well let's move on and talk about impact.

MEMBER RAMIAH: And the impact -- should I go onto the impact as the first point?

CHAIR BROTMAN: Sure. Let's start with impact.

MEMBER RAMIAH: Impact was the -- there was consensus that it was high impact in our subgroup and that the rates of these

STDs are higher in HIV population compared to the general population.

CHAIR BROTMAN: Any specific comments? Kathleen.

MEMBER BRADY: When do we have the discussion about whether it should be annual or once since diagnosis? Is that now? Is that under impact?

MS. WINKLER: Probably evidence more so than impact.

MEMBER GIORDANO: Just real quickly. The other impact is these sexually transmitted diseases also increase the rate of HIV transmission. So I think having them under control is believed to be an important prevention measure.

CHAIR BROTMAN: Thank you for that point. Okay, let's vote on high impact.

MS. KAHN: Voting on 1(a) high impact. You can go ahead and start. We have 11 high, 3 moderate, zero low and zero insufficient evidence.

CHAIR BROTMAN: Okay. So that overwhelmingly passes. Let's talk about the evidence now.

MEMBER RAMIAH: We discussed about the evidence presented was not specific to the STDs but to the general prevention efforts for the people living with HIV. And moreover, the measure was not aligned with the existing recommendation as mentioned, but the annual screening versus once, just once after HIV diagnosis.

It is also unclear as the -- how the screening can help with the performance improvement assuming that there is no sexual activity after one diagnosis. That was a gap in the evidence.

CHAIR BROTMAN: Any other specific evidence that anyone wants to discuss? You can bring up the preexisting point if you want.

MEMBER MILSTONE: Just a quick question. Was there any discussion in your

group about what to do with congenitally acquired HIV patients who are 13 who weren't yet sexually active? I'm thinking of how you could eliminate them. It would be really hard. I just didn't know if it was discussed.

MEMBER RAMIAH: No.

MR. REHM: I don't know if we were talking about the annual versus the -- our panel was literally split down the middle on this. And not vociferous for either side, but -- and I'll be frank. People who operated in the health plan environment -- Mary, you maybe can speak to this -- are concerned about overuse of a variety of services where you know, it's just the measure driving us to do something we know we don't need to do because we know Bob's and you know, whatever. You know, he just shouldn't be screened like that. And you know, trying to be respectful of that.

So, very seldom do we actually say we'll follow your lead but in truth our panel was split. We brought forward the one that

we came in with if you will. We can understand the utility. We also understand there's some unintended consequences of that as well. So, again, happy to get your input.

CHAIR BROTMAN: Peter? And then David.

MEMBER HAVENS: I do think that doing at least one screen can be looked at as an important improvement measure and would have impact since already there's many people don't get any screening at all. So, rather than get involved in a discussion that your own expert panel could not reach agreement on, might take this at face value and say it's worthwhile to do at least this. And if you want to come back with a potential second measure that would be more that could undergo testing or something else. But here this is as written an important measure for which there's a great deal of evidence if not just for prevention but also for routine screening in somebody who is universally, well, presumably sexually active.

CHAIR BROTMAN: David, go ahead.

MEMBER SPACH: I just was going to add it's possible that the measure could have been revised to just basically use some language similar to the STD guidelines, CDC STD guidelines that basically specify who needs recurrent testing. And that may have been one way around it.

CHAIR BROTMAN: Tom.

MEMBER GIORDANO: In terms of who should get recurrent testing or annual testing I think that is very difficult to operationalize and capture reliably. I'm very content with a measure that is sort of a minimum standard as long as there's evidence that people -- that we're currently not meeting the minimum standard.

And I would say that this is a minimum standard, screening everyone with HIV at least once for these important public health diseases. So in some ways I kind of, although we always present the evidence and then the

gap, I think in this case I want to see the gap data and then I'd say, okay, there's evidence that this is important. Is that possible? Can we see the gap data first?

MEMBER RAMIAH: Yes. The chlamydia and gonorrhea performance was 32.4 percent and syphilis was 50.3. was that right?

MEMBER GIORDANO: So for a single screen. Well, then I think there is room for improvement here which means this does have importance.

MEMBER HAVENS: Right. Without going to multiple screening or having a big argument. Exactly.

CHAIR BROTMAN: Well, let's first vote on the evidence. If there's no more discussion let's first vote on the evidence at this point.

MS. KAHN: Voting on 18 evidence. You can start your vote. You have 12 for yes, the body of evidence meets the guidance, 2 for no, the evidence does not meet the guidance

and 1 for insufficient information.

CHAIR BROTMAN: Okay. So that passes. And let's vote on the performance gap unless there's anything else to add.

MEMBER RAMIAH: No, nothing.

CHAIR BROTMAN: Okay. So let's vote on that.

MS. KAHN: Voting on 1(b) performance gap. You can go ahead and start. Can everyone press it one more time? You have seven high, eight moderate, zero low and zero insufficient.

CHAIR BROTMAN: Okay. So again that passes. Reliability.

MEMBER RAMIAH: The issue with reliability was mentioned earlier, the glitch, EHR glitch which caused 32-person difference between manual and -- manual inspection automated. For -- that was for chlamydia and gonorrhea whereas the syphilis one was only two-person difference. Right? So, that --

CHAIR SEPTIMUS: And this is the

one that there was a glitch in --

CHAIR BROTMAN: There was a glitch in the system.

CHAIR SEPTIMUS: -- EHR also.

MEMBER RAMIAH: EHR.

CHAIR BROTMAN: Yes, that they addressed as measure developers. Okay. Any discussion? Go ahead, Tom.

MEMBER GIORDANO: Are there data -- so after you fixed the glitch did it get better? Do you have that data?

MS. WILLIAMS-BADER: We weren't able to test that, no.

CHAIR BROTMAN: Okay.

MEMBER GIORDANO: These are laboratory -- the fact that it was done is captured in a laboratory. So I would think it would be reasonable.

CHAIR BROTMAN: You should be able to capture those. Kathleen.

MEMBER BRADY: But once again since it's since diagnosis I mean you could be looking

for data that is historically old and not in an EHR.

CHAIR BROTMAN: Okay. All right. Let's go ahead and vote for reliability at this point.

MS. KAHN: Voting on 2(a) reliability. You can go ahead and start. You have zero high, 10 moderate, 4 low and 1 insufficient.

CHAIR BROTMAN: Okay. So that passes. Let's go to validity.

MEMBER RAMIAH: The validity was done with the face validity and the number of N was 8 with a mean rating of 3.5. And it was a clear split between -- and it's mainly because of the comment between annual versus once right after diagnosis.

So, it's not a high face validity and the reasoning was that there was a discussion in the panel as if they should go for annual versus once after diagnosis.

CHAIR BROTMAN: Any specific

comments to that? Okay, let's vote for validity.

MS. KAHN: Voting on 2(b) validity. You can go ahead and start. You have zero high, nine moderate, six low and zero insufficient.

CHAIR BROTMAN: Okay. So that passes. On usability.

MEMBER RAMIAH: So, usability. It has been in use since 2010 and as Jenna mentioned has been in two different measures, one with chlamydia and gonorrhea together and syphilis separately and has been used in CMS PQRS with no issues to report.

CHAIR BROTMAN: Yes, go ahead, Aaron.

MEMBER MILSTONE: I just have a brief question and clarification. So the CPT procedure codes can get pulled out of the claims data, correct? So even if it was 7 or 8 years ago you could still pull it out of an old claims data, is that true?

MR. REHM: Recall that because the testing was done in the EHR environment we're talking about an e-specified measure in a way. CPT-2 is the PQRS program requirement. So remember this is true for all of our measures. Oh, I'm sorry, I thought I heard you say CPT-2. Excuse me.

MEMBER MILSTONE: No, no, this is the procedure code. I just didn't know if the procedure codes were through claims because then you could -- because we were discussing whether or not you're missing people who have transitioned from paper to electronic EHR.

MR. REHM: No. Yes, the CPT-2 -- pardon me, the CPT code is used in the health plan world where it's billing and we're receiving those bills. It's widely used.

CHAIR BROTMAN: Any other points? Let's vote on usability.

MS. KAHN: Voting on usability. You can go ahead and start. We have 2 high, 12 moderate, 1 low and zero insufficient.

CHAIR BROTMAN: And finally let's vote on suitability for endorsement.

MS. KAHN: We have to do feasibility.

CHAIR BROTMAN: Oh, feasibility. I'm sorry. Anything you wanted to bring up?

MEMBER RAMIAH: Nothing specifically. It's the same issues that came up in the usability also.

CHAIR BROTMAN: Okay. Any comments? All right, let's vote on feasibility.

MS. KAHN: Voting on feasibility. You can go ahead and start. Zero high, 14 moderate, zero low and 1 insufficient.

CHAIR BROTMAN: And finally now suitability for endorsement.

MS. KAHN: So does the measure meet NQF criteria for endorsement? You can start your vote.

CHAIR BROTMAN: Aaron, did you want to say something? That means we stop the vote.

CHAIR SEPTIMUS: Stop the vote.

MEMBER MILSTONE: Just one comment that might simplify this a little. So we were just saying you, in your denominator you actually restricted this to people who had two visits during the measurement year with 90 days in between. But we're looking at historical data, whether they've had ever had it. So I wonder why wouldn't this be anyone with a diagnosis of HIV that was seen in care during that year?

They should technically have had a test at some point, right? It doesn't matter whether they -- and that would make your data collection much easier. You don't have to restrict the denominator. It's anyone who has HIV that was -- had a visit. So I don't know if that -- just something to consider as you might simplify the measure.

MR. REHM: We're probably speaking in tongues here. You know, what's hard to appreciate is that remember when Jenna oriented

us to all these measures, some which have been endorsed and some which have not -- are suitable for endorsement.

This actually was the -- this, the office visit was a denominator so this two-visit to try to get at this balance between retention in care and people flying in, flying out, and understanding the tension between those two. So we maintain that denominator throughout all the measures which is why you're seeing it here.

And I think it was to have -- thinking from a suite perspective, consistency of that and that it met the single -- I understand your point and I think it's well taken. I think that that was the logic if you will.

CHAIR BROTMAN: Any comments on that? All right. Well let's finally vote for the suitability for endorsement.

MS. KAHN: You can go ahead and start your vote. You have 13 yes and 2 no.

CHAIR BROTMAN: Okay. So that passes.

CHAIR SEPTIMUS: Okay. We're going to put some time restraints on this next discussion. Okay, we're going to do about -- I apologize. My better half here, or better third.

The next thing on the agenda is the related and competing measures. We've had multiple discussions about that and the staff has put together some tables so we can see related to this and what we want to do about related measures. So I'll let Reva lead this.

MS. WINKLER: Yes. I mean, what has happened as a result of your discussions is most of this has become a non-issue. And so for all of the HIV measures that we identified potential related measures you didn't recommend both of them or all of them such that the only thing for HIV that still remains actually is two of the four visit measures. And the medical frequency, 2079,

the medical visit frequency and 2080 -- no, yes, that's right. You did not recommend 2081, the newly enrolled, or the 0403. So, you know, a lot of this issue about looking at them has fallen away.

The question is now there are two measures in this group that were recommended that do have sort of the focus around retention in care. And they are similar, they have different focuses if you will. Is there any question about the need for both? Or not? The first two.

2079, medical visit frequency, if you recall that's the HRSA measure that had a visit at least in each 6 months over a 24-month period. And then 2080 which is the gap in care and that was no visit in the previous 6 months of the measurement year. That's also from HRSA. So those are two that remain of this group.

MEMBER BLANK: Can I just ask a question for clarity on that? So, the -79 is

a different time period than the -80 in regard to what we're measuring but -80, lower is better. So you're looking for a lower percentage, a gap in care.

I was just asking for clarity for -79 and -80. There's a variation in the time period of measurement where -79, the higher the percentage is better wherein -80 it should be the lower the percentage is better. It's the reciprocal. The only variation to me was the time period in regard to how they were being measured.

CHAIR BROTMAN: They were approved.

MEMBER HAVENS: So what's your specific question for us that we could be the most helpful with right now?

MS. WINKLER: Again, we're looking at measures with similar -- that are very similar. The question is do we need more than one. Are they harmonized enough to be related and can work together. The fact that I think

the question was more pertinent when we had four. This was something that Ed particularly wanted us to be sure and do. But now that there's only two on the table perhaps it's less of an issue, but I want you to at least comment on it. Are you comfortable with having both of them?

CHAIR BROTMAN: Tom, did you want to comment?

MEMBER GIORDANO: Yes. I see them as complementary and not necessarily competing. I think that the first measure with 24 months requires that -- it's by definition a person who's been around for a little while in your clinic. And it's measuring their persistency with care, their retention in care.

The second measure is -- so that one will exclude new patients to the clinic because you have to have been in the clinic for at least 2 years. The second one can I think -- certainly will include new patients and will give you a slightly different look

at what's going on with your clinic population. It's a shorter measurement period which is in some ways beneficial. It's more inclusive of patients. And so that being said I think the first one's important too because it's a longer duration. So it's a different -- it is measuring something different. They are measuring something different. I would be in favor of both of them. I see no reason to try to force one or the other.

CHAIR BROTMAN: Thank you for that perspective. Any other perspectives in the room? Peter.

MEMBER HAVENS: I would just support that perspective. In a certain way 2080 is similar to 2081 in this kind of somebody who might have just joined and you fail to follow up. Frequently one of the problems with 2081 that seems like got it voted down was people couldn't agree on the number of visits that really were required for good care but we would agree that if you saw somebody once

early in the year and you didn't see them later that year that that seems like you have failed to bring them into care and keep them there, asks a different question than the long-term adherence to care which I think is -- which is 2079 and is very important.

CHAIR SEPTIMUS: Okay. Then I think that -- oh, Michael.

MEMBER FARBER: Yes, I think the thing on 2081 which many of the people, I think Peter as a matter of fact had stated was that the first measure gets you into the other three, but then the first measure which is that you got a visit in 4 months would be included in the other two which aren't the one who did the first one. So it could look like you did very poorly if you switched providers. So I thought that was a good point of why the 2081 was rejected.

MEMBER HAVENS: Well, and that goes to the central theme that went through all of these which I think is very difficult is that

as I understand the data and Tom, I would certainly defer to you on this, the provider type that is associated with staying in care and quality of care is an HIV-focused provider.

And from both HRSA and NCQA we heard about the difficulties in trying to identify that kind of provider type. Therefore we have to assume that these measurements are going to be applied to specific provider areas but NCQA was very clear that that is difficult to do in the context of an electronic health record. And I think that problem would pertain to the HRSA measure as well or measures. So I think that these are global questions of how to really look at what we're trying to look at. But I'm interested in Tom's take on that.

MEMBER GIORDANO: I think -- I don't disagree with anything you said but I do believe that -- and the data are more, are stronger for an HIV provider. That being said I think that as a minimum standard seeing anyone

is probably better than seeing no one. And so, and if I were a non-HIV provider and I had a patient with HIV who was coming to me for diabetes management or some other problem that I was okay managing I would be pushing that person to get to their HIV provider because I don't want to manage that and it's not my job to manage that but in the patient's interest I would try to push them there. So I think there is benefit. Even though it's not measured in the same way I think there is potential benefit to getting patients -- to keeping patients in any care.

CHAIR SEPTIMUS: Okay. I think the staff got the input they needed on these comparison measures. So if it's okay with you, Peter, I want to spend 10 or 15 minutes max on revisiting some additional information that you received on sepsis to give the developer fair hearing.

So just to review things. The vote on impact -- the impact was 19 high, evidence

was 11 high, opportunity was 7 high and 12 moderate. And where we got hung up was on the reliability issue because we didn't really have the data collection tool that was an attachment that apparently did not get received, did not get attached. So you all got that last evening on your way out.

No, no, we're finished with this. No -- well, they just wanted input whether or not these measures -- the input was that they're not the same and they're complementary. So, that discussion is finished. Tom, you look confused. I don't want to cut off discussion, Tom, but go ahead.

MEMBER FILE: -- unnecessarily putting two measures that are so similar that it causes confusion to the user or whatever. But I appreciate what you said, Tom. And I guess the real difference between the first and the second one is the second one would capture newly or newer patients in the first year, correct? So if they were seen once and

then weren't seen at the end of the 6 months that would be the difference, correct? And that you think has value. Yes, okay. Thanks.

CHAIR SEPTIMUS: Good, Tom. Now, we'll go back again with the sepsis. This is -- you got the data collection sample tool. There's also -- we didn't give you everything but there were several things that Manny sent earlier this morning, most of them related to the evidence and not to the reliability. There was some subsequent articles, two articles you sent later that went to reliability and data collection which is really where we hung up.

So this was a sample data collection that was supposed to be attached that you did not get on reviewing this before the vote yesterday. I also asked Helen and Reva to take this to their data folks to see what they felt about this data collection tool in terms of NQF's standards. So maybe I'll let either Reva or Helen address that point before we open this

up for general discussion.

MS. BURSTIN: Sure. So I had Karen Pace, our measure methodologist, review the testing that was submitted as part of this measure. And her overall perspective was what was submitted, granted it was a single institution, that the 498 charts, 9 reviewers in a single institution would pass our criteria for reliability. And that unless the committee had an a priori reason to assume that testing out of Henry Ford would not be representative of the rest of the nation it's not clear why that would have been an issue.

The bigger issue from our perspective is it was not clear yesterday how many people were voting on turning down the measure based on reliability based on the testing provided from Henry Ford versus the precision of the specifications which is specifically one of the elements of reliability. Since you didn't have the detailed data collection tool which was our

fault, we just felt like we had a process misstep that we just felt like we needed to go back and have you re-look at that.

In addition, Dr. Rivers sent along additional information this afternoon but I think he also wanted to convey, and we're just calling him to have him dial in, that, you know, about 100 hospitals currently who are using the measure, all of them have an internal audit process that always looked to see a sample of the charts to see if the data for the bundle is reliably collected.

So he personally spoke with Kaiser, Sutter and a couple of other health systems overnight and they confirmed they all do an internal audit. This is very analogous to what tends to happen with our registry-based measures like STS and ACC where there is a sub-sample of measures that institutions review on an audit trail to see if they're appropriately being collected.

So we just wanted to bring it back

to you. If you feel like, you know, this is the right time to do it that's fine. We could give you more time but I defer to the chairs here.

CHAIR SEPTIMUS: Sure. Thomas?

MS. BURSTIN: Dr. Rivers is on with us as well, by the way.

CHAIR SEPTIMUS: Okay. Well, let's Tom -- go ahead, Tom.

MEMBER FILE: I appreciate this. And I'm sympathetic to the fact that this was not here because that was one of my concerns was the precision of the data, you know, and why I was concerned about it. I still am concerned about it but I'm less concerned I guess.

But -- and maybe I shouldn't say this but I think it's a little disingenuous at the last hour to give us this. I mean this is based on the evidence which we already agreed on. I mean, there was a big consensus that there was good evidence so I don't think we

need that.

But I'm still concerned about the reliability. I'm glad at least we have this because when I'm looking at the criteria for precision of specification and repeatability and now you give me the data from Henry Ford. What I have to know, were the data extractors from Henry Ford, were they part of a research team or is that a total independent, untrained, not -- I shouldn't say untrained because data extractors are trained -- but not part of the clinical trial who have obviously a different knowledge base than an independent data extractor would have.

Because when I look at this now very quickly and I looked at it last night is this exactly what Henry Ford did or what they're -- because it looks like there's examples of data collection here. One says Cooper University Hospital so I don't know where that comes from. The other looks like it's from Surviving Sepsis campaign which to me looks

more like a tool for data extraction for a database.

DR. RIVERS: This is Dr. Rivers.

CHAIR SEPTIMUS: Let Tom finish and then we'd like you to respond to his comments.

DR. RIVERS: Oh, okay. I'm sorry. I just called in.

MEMBER FILE: And please correct me because I want to vote for this. But, and that's why when you said, Helen, that your data extractors say that this meets the standard, if it really meets the standard then I'm going to vote for it. But I just want to make sure it's clear that -- because when I look at specifications.

For example, on one of them it says -- and I apologize, I'm probably hung up on this because I do so much antibiotic stewardship and I want to make sure the appropriate antibiotics are used. The check is was broad spectrum antibiotics given. Now, who interprets that? I mean, there's just a

checkbox.

Now, when we do other measures for you have to use -- well, it's like when we were talking about antiretroviral therapy and they said you have to use antiretroviral recommendations that are in the most recent guidelines. Well, at least then we have a source that we can say, well, were these regimens used. I don't know, where's the specification of what antibiotics can satisfy this measure? That element of the measure. So that's my point. And I'd like to vote for this, I just have to be convinced.

CHAIR BROTMAN: Dr. Rivers, can you respond to any of that?

DR. RIVERS: Oh, sure. And I perfectly understand it. The other -- most important is the first antibiotic must be in within the first 3 hours and the basis for that is most antibiotics broad spectrum will cover pretty much 90 percent of the bugs.

Now, the IDDS have a recommendation

for antibiotic regimens based on empiric antibiotics based on location of infection and many of these guidelines are based on that. So the key point is that that is the first antibiotic choice. So if you left it up to a clinician and these are multiple studies that have looked at antibiotic correctness after just empirically giving one dose or based on the clinician's suspicion of where the site of infection is, they're correct 90 percent of the time when those cultures come back. So with that background the key point is just get the antibiotic in and no matter what the antibiotic choice is don't get hung up on antibiotic choice because it's usually 90 percent of the time it's correct. It's just get that dose in.

And then the infectious disease gets involved and perhaps maybe modify that antibiotic. But that first dose, this is what is based on that first dose.

CHAIR BROTMAN: Tom, go ahead.

MEMBER FILE: Then I agree with that. If you just took away -- if you said antibiotic given within the first hour or first 3 hours or whatever, fine. I would totally agree with that.

DR. RIVERS: And that's what -- that's all it is. It's not to look at correctness or anything because again that's based on cultures but that takes time for those to come back.

CHAIR SEPTIMUS: Okay. Any -- so Tom, you'd be okay if we took out that broad spectrum. Just antibiotics administered within a certain period of time.

MEMBER FILE: I still have to be convinced that this document -- but if you're saying this document satisfies the standard.

MS. BURSTIN: I'm not talking about the document. What we shared with Karen who was on the phone earlier today was the actual testing submitted by Henry Ford in the submission form. The 498 --

MEMBER FILE: Okay. Well, I need the --

MS. BURSTIN: -- yes.

MEMBER FILE: -- question answered as well.

MS. BURSTIN: Yes.

MEMBER FILE: Who did that testing? It was non-clinicians?

MEMBER BRADY: It was clinicians.

MEMBER FILE: I mean, is that who's going to be doing the data extraction for the charts for all the charts in the whole measure?

CHAIR SEPTIMUS: Dr. Rivers, can you answer that?

MEMBER FILE: Are you going to require clinicians to do all the -- in our hospital, for example.

DR. RIVERS: Yes, the preferable solution would be to have a -- what they call, we have a sepsis nurse who basically is responsible for capturing all patients as well

as the database. So at minimum you want somebody who's familiar with each one of those variables and familiar with all of the nuances of data capture. Specialty, it doesn't matter, but it has to be in most places a clinical nurse. Or either somebody who's been in the clinical arena for an experienced period of time.

CHAIR SEPTIMUS: And for HCA in our 55 hospitals that are now engaged in this it is a sepsis coordinator that enters the data in our database. It sounds similar to what Henry Ford does. Mary?

MEMBER BLANK: I was just going to comment. In my experience we have this as one of our pay-for-performance initiatives. It is the quality department that abstracts the data. So not even a coordinated sepsis nurse but the criteria is listed for each of the metrics.

CHAIR SEPTIMUS: Here comes Aaron.

MEMBER MILSTONE: So just trying

to clarify what our intention is. Are we going back and re-discuss this from the beginning? Because there's other additional information. Like I know Tiffany brought up one of the questions about some pushback on including CBP monitoring. And I'm looking at the tables that were provided also. There are a number of studies here that don't show significance in recording CBP as one of their covariates in multivariable analysis.

So I think if we're going to -- I'm just trying to gauge are we -- is the intent that we're going to re-vote on this or just re-discuss? Because if we're going to re-discuss I wonder whether -- with new data I wonder whether we need to re-discuss with new data.

MS. BURSTIN: That would be up to you. I mean, at this point I think our feeling was we didn't give you adequate information to assess reliability. We viewed that as a process issue. We just wanted to correct that.

We don't necessarily feel the need to go back to evidence unless you do. You had a quite extensive discussion on evidence yesterday. So again, I think what we'd like to do is, and I'm now getting emails from somebody at Sutter Health providing additional data as well. So it's fast and furious here.

I think we would just want to be as fair as possible. If you feel like this is too much to digest at the eleventh hour here we can also try to package it, put it forward to you. It's always harder to do these things after the meeting, that's all. So I defer to Ed and Steve on that.

CHAIR SEPTIMUS: The evidence was not where we got hung up. Okay? And we also discussed bundle versus single elements. We went through all that. And I, unless everybody else -- I think we got hung up on the reliability and validity of the data, not on the impact or the evidence of the measure.

CHAIR BROTMAN: And I thought we

did have a relevant discussion, a robust discussion on the CBP and a couple other issues at the time. But again, want to hear from you if that's necessary.

CHAIR SEPTIMUS: The other option is to digest this and take this up at another time. So, again, I think everyone sort of feels somewhat bad because part of the document was not given to you ahead of time. It wasn't given to you after we voted on it yesterday. And so we think that was -- that would have been relevant to the discussion.

MS. BURSTIN: It also sounds like there might be additional data from the Surviving Sepsis campaign database that we could bring to bear. This is the note I just got from Dr. Townsend at Sutter Health. So again, if you guys would rather have us package this cleaner, get it out to you and have that discussion offline we can do that. I just wanted to at least bring it up because I think again from our point we've got to be really

careful about process and I don't think we met it yesterday, that's all.

CHAIR SEPTIMUS: Peter.

MEMBER HAVENS: So, one of the central differences between this bundle and the CLABSI, the central line insertion bundle is that this bundle includes an invasive procedure and excludes people who didn't have the invasive procedure.

And the denominator problem leads to a difference in who you can apply this to. And the need to have a central line may -- well (a) you can't then evaluate that if you're only looking at people who got a central line, and (b) in the Journal of Intensive Care Medicine paper that was just supplied to us, published online 17 August 2012, central venous pressure achieved is not statistically significantly associated with outcome in Table 4.

And while the central venous oxygen saturation greater than 70 percent was

statistically significant at P of 0.047 it's not clear that that was achieved because of use of the central venous pressure monitoring or because of an administration of a blood transfusion which also acts to bring up oxygen delivery and therefore would increase central venous measured oxygen saturation.

So, the bundles as we discussed earlier are markers of hospital systems activity on the one hand and also may have components that are more or less important to the outcome of the patient who is cared for in a bundling of services. And it's one thing to approve a bundle without an invasive procedure, but a completely different problem to approve a bundle that includes an invasive procedure and excludes people who don't get that procedure.

CHAIR SEPTIMUS: I think the co-chairs are going to make a decision. We are losing people. It's towards the end of the hour. If it's okay with the committee we'd

like to carry on this discussion online but I think it would be unfair given the late hour and given the complexity of the discussions. Let's get all our ducks lined up in a row and get this information out to you in a format that I think is meaningful. I think we can finish the discussion at another time where we have appropriate focus on it. Is that agreeable to everyone? I think we're going to -- no matter which way we vote we may be doing a disservice to the measure either up or down.

MEMBER BRADY: I would add some of the experts specific to this particular indicator have now left.

CHAIR SEPTIMUS: Excuse me?

DR. RIVERS: I'm still here.

CHAIR SEPTIMUS: Okay. So, Manny, we're going to postpone the completion of this because people are leaving and we're going to bring it back in another format online.

Is there any, Operator, in the room,

anyone for public comment before we adjourn?

OPERATOR: If you would like to ask a question please press *1 on your telephone keypad.

MS. WINKLER: No. Just if we don't have any, one other thing. We talked about disparities throughout the day. One of the things Nicole and I have been doing all along has been looking to see how your comments feed into our disparities protocol. And so what we're going to do is provide you sort of a conclusion of how we are viewing the measures from a disparities-sensitive perspective for you to comment on. And we'll give that to you offline and let you comment. So that's that.

The other thing is again since we're going to be chatting virtually one thing we always ask all committees is, okay, these are the measures you had in front of you for the topic area of infectious disease. Was there anything glaringly missing? I mean, are there really important aspects of care for which

there aren't any measures that you would recommend that measured development be pursued?

CHAIR SEPTIMUS: Antimicrobial stewardship is a big void that many of us have talked about offline. And there is some discussion with several to do that, that one of the big glaring gaps in ID is antimicrobial stewardship. Kathleen?

MS. WINKLER: And like I say, this is something -- since we're going to be chatting a lot feel free to forward your suggestions. But that is always something, given that you've spent so much time looking at the measures that are, perhaps you have some thoughts on the measures that should be and are not.

CHAIR SEPTIMUS: Kathleen?

MEMBER BRADY: And I mentioned this to Reva --

DR. RIVERS: This is Dr. Rivers.

CHAIR SEPTIMUS: Yes, Manny.

DR. RIVERS: There is a big statement coming out, a consensus for procalcitonin use in infectious disease. And that -- AHRQ through their -- that's going to be published soon and may be a good idea. There are many collections throughout in terms of the use and implications of procalcitonin. That may be something to look at.

CHAIR SEPTIMUS: Yes, sort of a parallel to stewardship, but yes. And Kathleen?

MEMBER BRADY: And I mentioned this to Reva earlier but it -- about HIV testing in persons ages 13 to 64. And I think that actually belongs in the infectious --

CHAIR SEPTIMUS: So I have passed that. I'm okay then.

(Laughter)

MS. WINKLER: We get to evaluate you based on risk.

(Laughter)

CHAIR SEPTIMUS: Steve I'm sure

will have -- it's been really an honor to be asked to co-chair. This is an incredible amount of talent around the room. I know that I certainly learned an enormous amount over the last day and a half, almost three-quarters of the day, and I hope that we'll continue to learn from each other. And I thank you for your attention. And I'll let Steve make the final comments.

CHAIR BROTMAN: I just want to thank everyone for bringing their brain trust to the table. And we'll have continued conversations but it's been nice meeting everyone in person. So, safe travels. Thank you.

(Whereupon, the above-entitled matter went off the record at 3:25 p.m.)