NATIONAL QUALITY FORUM

+ + + + +

CARDIOVASCULAR MEASURE ENDORSEMENT PROJECT
STANDING COMMITTEE MEETING

+ + + + +

MONDAY
APRIL 21, 2014

+ + + + +

The Committee met at the National Quality Forum, 9th Floor Conference Room, 1030 15th Street, N.W., Washington, D.C., at 11:00 a.m., Mary George and Thomas Kottke, Co-Chairs, presiding.

PRESENT:
MARY GEORGE, MD, MSPH, FACS, FAHA (Co-Chair), Centers for Disease Control and Prevention, Division for Heart Disease and Stroke Prevention
THOMAS KOTTKE, MD, MSPH (Co-Chair), Medical Director for Population Health, Consulting Cardiologist, HealthPartners
SANA AL-KHATIB, MD, MHS, Duke University Medical Center
LINDA BRIGGS, DNP, George Washington University, School of Nursing
JEFFREY BURTON, RN, Clinical Performance Improvement Specialist, United Physicians
LESLIE CHO, MD, Cleveland Clinic
JOSEPH CLEVELAND, MD, University of Colorado Denver
MICHAEL CROUCH, MD, MSPH, FAAFP, Texas A&M

University School of Medicine
ELIZABETH DeLONG, PhD, Duke University Medical Center
TED GIBBONS, MD FACC FACP FASE, Harborview Medical Center; University of Washington School of Medicine*
ELLEN HILLEGASS, PT, EdD, CCS, FAACVPR, FAPTA, American Physical Therapy Association
JUDD HOLLANDER, MD, FACEP, The University of Pennsylvania
THOMAS JAMES, MD, AmeriHealth Caritas Family of Companies
JOEL MARRS, PharmD, FNLA, BCPS (AQ Cardiology), CLS, Skaggs School of Pharmacy and Pharmaceutical Sciences, University of Colorado Anschutz Medical Campus; American Society of Health-System Pharmacists
KRISTI MITCHELL, MPH, Senior Vice President, Avalere Health, LLC
GEORGE PHILIPPIDES, MD, Boston University/Boston Medical Center
NICHOLAS RUGGIERO, II, MD, FACP, FACC, FSCAI, FSVM, FCPP, Thomas Jefferson University Hospital
JASON SPANGLER, MD, MPH, FACPM, Amgen, Inc.
CHRISTINE STEARNS, JD, MS, NJ Business & Industry Association
HENRY TING, MD, MBA, Mayo Clinic
MARK VALENTINE, MBA, The Heart Hospital Baylor Plano, Baylor Health Care System
MLADEN VIDOVICH, MD, Jesse Brown VA Medical Center

NQF STAFF:
HELEN BURSTIN, MD, MPH, Senior Vice President, Performance Measurement
WUNMI ISIJOLA, MPH, Project Manager
VY LUONG, Project Analyst
LINDSEY TIGHE, Senior Project Manager
REVA WINKLER, MD, MPH, Senior Director

ALSO PRESENT:

KYLE CAMPBELL, PharmD, FMQAI*

JENSEN CHIU, MHA, American College of Cardiology

FRED MASOUDI, MD, MSPH, FACC, American College of Cardiology

SOEREN MATTKE, DSc, MPH, RAND*

BRAHMAJEE NALLAMOTHU, MD, American College of Cardiology

LARA SLATTERY, American College of Cardiology

*** present by teleconference**

A-G-E-N-D-A

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

11:05 a.m.

MS. ISIJOLA: Good morning, everyone, and welcome to the cardiovascular standing committee. It's really great to put some faces to some of the names that we've been working with over the past few weeks.

My name is Wunmi Isijola. I'm the project manager here at NQF. And I just want to kind of give you an overview of what we're doing today just of our agenda.

So, first, we're going to start off with some introductions. And I will turn it over to our general counsel during that time to talk about the disclosure of interest followed by some of the roles and responsibilities as you our standing committee members.

Then we're going to follow off with our portfolio review of the cardiovascular measures. And then we'll get started with consideration of the candidate

measures that we have within this project. And I just wanted to turn it over to Helen.

DR. BURSTIN: Good morning, everybody. Just to add my welcome. Helen Burstin. I know many of you and thank you for coming back for those of you who worked with us the last round.

We're excited. This is one of our first standing committee meetings. We had one last week as well. And just the idea of having a group who has that knowledge over time and can bring measure issues back to you.

We did an ad hoc review, for example, as part of our safety measures project. It was really a huge advancement for us around harmonization, alignment, keeping up with the science, keeping up with the evidence. So I'll be glad to join you for this and thanks, all, for coming.

MS. ISIJOLA: Thank you, Helen. And I just wanted to introduce our staff here at NQF.

Again, my name is Wunmi Isijola. I wanted to also introduce you to Vy Luong. Many of you have been in contact with her over the past few weeks. And we have our senior project manager Lindsey Tighe who's here as well. And we have Dr. Reva Winkler, our senior director on the project.

And I also wanted to briefly introduce our co-chairs. Oh, there's Vy. Say hi, everyone. I also want to introduce our co-chairs who will really be facilitating the discussion today.

We have Dr. Thomas Kottke and we have Dr. Mary George. We do appreciate, again, everyone being here and really your efforts over the last few weeks.

And with that being said I will turn it over to Ann Hammerstein for our disclosure of interest.

MS. HAMMERSMITH: Good morning, everyone. I'm Ann Hammersmith, NQF's general counsel. We're going to combine the

introductions with the disclosures because it's a bit quicker that way and we want you to be able to get to your work.

I see a few familiar faces so some of you have heard what I'm going to say before. But I will say it again.

Just a few reminders. You received a form from us to fill out where we asked you about your professional activities and so on which you turned in and we reviewed.

What we like to do at the beginning of the meeting, the first meeting, is to have you go around the table and disclose anything that you think is relevant based upon that form and based upon your activities.

I want to remind you you sit as an individual. You do not sit as a representative of your employer. You do not sit as a representative of anyone who may have nominated you to serve on the committee. You're here because you are an expert.

Unlike a lot of conflict of interest processes we look at things other than financial issues. So, if you have served on a committee, you may have served on a committee as a volunteer. And if it is relevant to the work that the committee will do then we would look for you to disclose that.

In addition, I want to remind you that just because you disclose doesn't mean you have a conflict. Part of the point of this exercise is for people to understand where everyone is coming from and what their background is.

We do ask you not to summarize your resume, please. Only disclose things that are relevant to the committee's work. We are particularly interested in any grants, research or consulting work that you may have done, but only if it is relevant to what the committee will be looking at.

So with that I'll start with the

chairs. I always make the chairs go first.

DR. GEORGE: Good morning and welcome. I'm Mary George from CDC in Atlanta. In terms of conflict of interest I was on the previous cardiovascular steering committee and also on an ad hoc NQF committee that reviewed some updated risk-adjusted mortality measures for heart disease.

Other than that I don't have any other conflicts of interest.

DR. KOTTKE: Tom Kottke from HealthPartners in Minneapolis, in St. Paul. I was on the prior cardiovascular committee. Otherwise no conflicts of interest.

MS. STEARNS: Christine Stearns with the New Jersey Business and Industry Association. I work for a trade association with 21,000 businesses in New Jersey. This is the third actual NQF panel that I've worked with. I was on the previous Cardiovascular Steering Committee.

DR. HOLLANDER: Judd Hollander.

I'm an emergency physician at the University of Pennsylvania. But by our last conference call I'll be an ER doc at Jefferson.

I don't believe I have any direct conflicts of interest related to the measures that we're reviewing today.

DR. CLEVELAND: Good morning. I'm Joe Cleveland. I'm an adult cardiac surgeon at the University of Colorado here I guess representing -- not representing, but nominated by the Society of Thoracic Surgeons.

So in that realm I do have some disclosures. I do and have served on the Quality Measurement Task Force for the STS. And also and currently a member of the Adult Cardiac Surgical Database for the Society of Thoracic Surgeons. Those would be the only disclosures that I think I have. Thank you.

DR. JAMES: Good morning. Tom James. And I'm not related to the Tom James who did all the electrophysiologic work.

I'm the medical director for

clinical policy at AmeriHealth Caritas, a managed Medicaid company. I co-chair the AQA Public Reporting Workgroup and chair the NQF Health Plan Council. And those are my only disclosures.

MS. HILLEGASS: Ellen Hillegass. And I'm a representative or referred by the American Physical Therapy Association. I'm an APTA board certified cardiovascular and pulmonary specialist.

And I really don't have any disclosures except I was recently appointed to it's called Quality Insights Task Force which does have a measure that was pulled for tomorrow. And I did make you all aware of that.

DR. VIDOVICH: Mladen Vidovich. I'm from University of Illinois-Chicago and I'm chief of cardiology at Jesse Brown VA in Chicago. I don't have any direct conflict of interest related to this.

I was also recently appointed as

the governor-elect for the Department of Veterans Affairs at the American College of Cardiology.

MS. DELONG: Liz DeLong. I'm at Duke University. I confess that I have worked on several cardiovascular databases including the NCDR and the STS. I've also served on some previous NQF committees not specifically related to cardiovascular.

DR. RUGGIERO: I'm Nick Ruggiero, interventional cardiologist at Thomas Jefferson University in Philadelphia. And I have no conflicts of interest.

MS. BRIGGS: Hi, I'm Linda Briggs. I'm a nurse practitioner and faculty at George Washington University School of Nursing. My background is cardiovascular and I've worked both medical and surgical cardiology. I have on conflicts and no disclosures.

MR. VALENTINE: Hello, I'm Mark Valentine. I'm the president of the Heart Hospital Baylor Plano and the Heart Hospital

Baylor Denton. I've been an administrator for the last 23 years. I have no conflicts.

DR. CROUCH: I'm Michael Crouch. I'm a family physician at the Memorial Family Medicine Residency in Sugarland, Texas. I served on a previous cardiovascular NQF committee. I have no other conflicts of interest.

MR. MARRS: Hi, I'm Joel Marrs. I'm a clinical pharmacist and a faculty member at the University of Colorado. And no conflicts of interest to disclose.

DR. SPANGLER: Good morning, I'm Jason Spangler. I'm executive director of medical policy at Amgen. And as part of that role I lead our quality strategy for the company. So, I am an employee of Amgen.

We don't have any current cardiovascular products but we do have a couple of products in the pipeline. None that are directly related to the work that we're doing for this right now in the next phase

possibly which I will disclose at that time. Thanks.

MS. MITCHELL: Good morning, my name is Kristi Mitchell. I am a senior vice president at Avalere Health. I have no conflicts to disclose. However -- rather no conflicts of interest, but rather I spent 12 years at the American College of Cardiology. So it's kind of hard to put that in a box.

And as a result I led the development of the National Cardiovascular Data Registry and several of the measures that we are talking about as a staff member.

DR. CHO: Hi, I'm Leslie Cho. I'm an interventional cardiologist from Cleveland Clinic and I head the preventive and rehabilitation section. I've served on previous NQF committees and I currently serve on the technical expert committee.

DR. PHILIPPIDES: Good morning, my name is George Philippides. I'm a cardiologist at Boston University Medical

Center. I was on the prior cardiovascular committee for NQF and have no disclosures. Thank you.

DR. AL-KHATIB: Good morning, I'm Sana Al-Khatib. I'm a cardiac electrophysiologist at Duke University. And I don't have any conflicts of interest in relation to the measures that we will be discussing today. But I have worked on performance measures. I co-chair the Measure Development Task Force for the heart rhythm society. I am on the steering committee for the NCDR ICD registry. And I am working -- with a working group to inform the development of performance measures for ACC.

DR. TING: Good morning, I'm Henry Ting. I'm a cardiologist and health services researcher from Mayo Clinic.

I do have some conflicts which I've disclosed. I'm on the ABIM council. I also am on the American College of Cardiology and American Heart Association Task Force for

Performance Measures. And I've participated in grant work and received grants from AHRQ and NHLBI to develop some of these measures which I've disclosed and recused myself from.

MS. HAMMERSMITH: Okay, I'm going to call on two committee members who are on the phone to disclose. Ted Gibbons?

DR. GIBBONS: Hi, I'm Ted Gibbons. I'm at the University of Washington and I'm chief of cardiology at Harborview Medical Center, the public health hospital associated with the University of Washington.

I have been on previous NQF cardiovascular committees and I have nothing else to disclose.

MS. HAMMERSMITH: Okay, thank you. Jeff Burton? Is Jeff Burton on the line? Are there any other committee members on the line? Okay. Thank you for making those disclosures.

And my parting words to you are to make sure that you understand that you are important parts of a successful disclosure of

interest process and policy.

If you are sitting in a meeting and you think you may have a conflict, if you think a fellow committee member may have a conflict, or if you think someone is behaving in a biased manner please do speak up.

We don't want you sitting there in silence if you think that there may be a conflict. You are welcome to bring anything up openly in a meeting. You can go to your co-chairs who will consult with the NQF staff, or you can go directly to NQF staff.

So in that spirit do you have anything that you wish to discuss with each other, or do you have any questions of me based upon the disclosures made this morning? Okay, thank you.

DR. SPANGLER: I didn't think about this probably before because I didn't think it was relevant but I want to bring it up just in case there's a question.

We do have a product that was

launched in Europe that's going to be launched here in the U.S. It's a heart failure medicine. And one of the measures I was the primary on was a heart failure one. And it had to do with a follow-up appointment. So, I didn't think it was directly relevant but I wanted to throw that out there in case people thought it was.

MS. HAMMERSMITH: Okay. I would say it isn't. You know, if the measure directly implicated that class of bugs then yes, we would have something to talk about.

DR. SPANGLER: That's the thought I had but I wanted to throw it out there.

MS. HAMMERSMITH: Yes.

DR. HOLLANDER: Just because everybody else mentioned committee work, I'm on the Quality and Performance Committee for the American College of Emergency Physicians and they're the people that nominated me. I don't think it's a conflict.

MS. HAMMERSMITH: Okay. Thank you

for disclosing it. Anything else? All right, thank you.

MS. ISIJOLA: Thank you, Ann. Okay. So, I just want to go over some of the roles of standing committee members. I know we went over this during our orientation, but just to kind of give you a recap of some of the expectations that we expect from you as standing committee members.

Given your backgrounds we do expect you to act as a proxy for our NQF multi-stakeholder membership and serve on at least a 2- or 3-year term. We will be designating your term assignment during day 2, randomly selecting that. But if you do have any concerns with that please do let us know at that time.

And essentially we ask that you work with our staff to really achieve the goals. And in this case it's really evaluating the measures that are at hand. And evaluating this based on the criteria that was

presented to you which helps to indicate the extent to which each criterion is met as well as the rationale for the rating.

We also ask that you make recommendations to the NQF membership for endorsement by essentially responding to any comments submitted during that review period, but also responding to any direction given by our CSAC committee.

And lastly, really just overseeing the portfolio of cardiovascular measures in which we have roughly about 80 measures at this point. So, these are kind of some of the expectations.

And really the reason why we opted out to really hold the standing committee is because we wanted to ensure that there is consistency across the board. I mean, once you're starting to look at these measures you get a sense of what's in our portfolio, what are some of the gaps. So we ask based on your expertise to provide that input.

And knowing which measures are in our portfolio and understanding what the importance of these measures are.

And like I mentioned, just really identifying what are the gaps within our portfolio. Because I know we do hit on some of the subtopics but there are a wide range of measures that aren't necessarily in our portfolio today.

And we ask that you are aware of the measurement activities for this topic area and know there are up and coming guidelines within the cardiovascular arena. So we ask that you are cognizant of that and really bring your input as we look through this portfolio.

And lastly, just providing your feedback about the evolution of our portfolio and considering additional new measures in which you would like to see within the cardiovascular topic area.

So, today we have 17 measures that

we will be reviewing over the next two days. And these are just a snapshot of what those measures are. I know we did send an email to you guys about a measure that was withdrawn at the last minute so we will be reviewing 17 of those measures within this project.

And now I will turn it over to Dr. Winkler and she will kind of give you an overview of what our portfolio looks like and what measures and how that kind of translate into what we're looking at over the next two days. So, Dr. Winkler?

DR. WINKLER: Thank you, Wunmi. I'd like to turn the committee's attention to your Sharepoint site because one of the committee documents we have provided for you is an overview of the portfolio.

The cardiovascular portfolio for NQF is one of our largest and it does encompass a wide range of topic areas and measures. One of the things that makes it a little bit easier to get your arms around is

to figure out an appropriate framework for organizing and presenting the measures within the framework. And so I do want to kind of go through how we've organized them. But we're certainly open to any input from you in terms of how we might want to improve the framework of the organization around cardiovascular measures.

To start off, the cardiovascular topic area is really a very important one. One of the measure priorities from the National Quality Strategy which NQF tries to work with the National Quality Strategy in all the work that we do. One of the priorities is promoting the most effective prevention and treatment practices for leading causes of mortality. And this is where we come in, starting with cardiovascular disease.

So, this is sort of the first topic. It's by no means the only important topic in NQF's portfolio but certainly it is a high-profile one for the NQS. So keep that

in mind as we're looking at this portfolio.

Now, we have a lot of different topic areas. We currently have more than 70 endorsed measures that are in this portfolio that you actually oversee.

However, in some of our other topic areas portfolio there are related measures. So we want you to be aware of them and not look just at your particular group in a vacuum but understand that there is crossover. And sometimes our assignment of measures to portfolios is a bit arbitrary.

And so being aware that there are other measures out there that may be related helps you understand how any measure you're evaluating fits within not just the cardiovascular portfolio but NQF's portfolio of measures totally.

So we have measures around coronary artery disease and acute myocardial infarction. It's one of our biggest subsets of measures. So we're going to take a look at

those.

We also have a goodly number of measures around heart failure, around rhythm disorders, cardiac cath, actually very few around hypertension considering its importance overall and then some cost and resource use measures.

So, organizing this group was a bit of a challenge. But NQF has been working in this area for quite awhile.

And so a couple of years ago, it's now -- it's getting on five or six years ago now there was a project in which a group of folks were looking at patient-focused episodes of care. And actually they created a patient-focused episode of care diagram to really look at it from the patient's perspective of focused on the acute myocardial infarction as the episode, but of course realizing there are a lot of antecedent events, there are a lot of related events that could occur.

So if you take a look at this,

I've heard these referred to as NQF's bubble diagrams. And so you can see that there is a large population at risk, either primary prevention perhaps, certainly secondary prevention in patients that have exhibited their coronary artery disease. So as that large underlying population.

There may be an acute phase, an acute event such as an AMI but perhaps it's more of a procedure such as a PCI or a CABG or some other acute event for which there may be an acute phase and care organized around that acute event.

After an acute event there are post-acute care, rehabilitation phases. And then those folks again sort of circle back into the secondary prevention.

The sense was that patients follow several different potential trajectories. Some somewhat more stable and progress on to a relatively stable situation where focusing in on maintaining functions, secondary

prevention is really important.

Another pathway of course are those with significant cardiac damage and issues around quality of life, advanced care planning, palliative care, may be more appropriate.

So as we look at this sort of spectrum of care we organize the measures for coronary artery disease and acute myocardial infarction according to these different bubbles or phases because it seems to reflect the patient experience. And again, your input into this would be perfectly welcome.

And so honestly we do have measures in all the bubbles. The question I think given we've got a large number of measures is do we have the right measures. Do we have measures -- do we have an efficient number of measures. And so again, as part of your oversight this is the kind of input and conversation we'd like you to have.

So, if we look at the measures

around a population at risk for primary prevention you'll see that we do have several measures around smoking prevalence, tobacco use screening, some cardiovascular screening in certain populations, blood pressure screening and control.

Some of these are not in this particular cardiovascular portfolio, particularly the tobacco measures because those are in our what we call health and well-being portfolio because they apply across the board.

MS. DELONG: Do we have these? Are we supposed to be following you here?

DR. WINKLER: There is a document in your Sharepoint. I don't think it's necessary today right now, but I think after we've had a chance to talk you may find it useful to refer to.

So the population at risk, the primary prevention. Also we have several measures around cardiac imaging, stress

imaging for relatively low-risk patients, non cardiac low-risk patients, pre-operative evaluation. Again, looking for the patient I think that may have previously undiagnosed cardiac risk factors.

So, again, a fairly substantial number. Perhaps there are other measures that are gap areas that you could consider. But again, as we have our conversations going forward not just today but as the committee is looking at measures you may want to think about measures that would be more appropriate gap areas, or where do we move on from here. So these are sort of the first bubble if you will.

We talk about secondary prevention. And this is another large area of measures around blood pressure management, antiplatelet therapy, ACE inhibitors, lipid control, blood pressure control, so all the usual characters.

I will say that we have -- are

postponing maintenance review of any of the measures having to do with blood pressure and lipid control in this immediate time frame because of the recent new guidelines. We're giving developers time to adjust measures to those new guidelines. So it's not that we're not interested, but we are spending a little bit -- giving them a little bit of time to adapt to the new guidelines. So these measures are in the portfolio and you will be seeing them in the next couple of years.

So, the next group is again acute phase. I think these are probably measures well known to everyone. They are hospital-level measures as well as clinician-level measures for the care of patients with AMI. Again, many of them are -- the hospital-level measures are reported on Hospital Compare. They've been reported for a long time. We're certainly seeing some high levels of performance at this point in time.

And so you'll see that for this in

red I've put in the measures that are newly submitted. So they are not part of the portfolio yet. But this is a measure that you will be evaluating tomorrow. And you can see that it's a composite measure. And I'll be very interested to hear your reaction to a composite measure given the number of other measures that already exist in this area.

In the acute phase around AMI we have outcome measures as well as the number of process measures. I think you're all aware of those. Again, many of them reported on Hospital Compare.

Certainly the readmission measures. We are looking at all readmission measures together in another project. So it will not come to you at this point in time. The readmissions measures are being evaluated by a separate committee. They're meeting next month. So there will be sort of concurrent discussion around the AMI readmission measure that you might be interested in.

We also have mortality measures that are inpatient as well as sort of the 30-day all-cause measure that I think you're familiar with.

One of the new measures for tomorrow. Again, leading edge, not unusual in the cardiovascular portfolio is a 30-day mortality eMeasure.

So this is one of the first, in fact I think it is the first eMeasure that's an outcome measure. We have a couple of that are process measures but eMeasures are sort of a new and up and coming thing. And I think there is the hope that as we're able to transition measures to use the unique characteristics of EHRs eMeasures will become an important aspect of the portfolio. So, you get the first one. So we'll be talking about that measure tomorrow.

So outcomes are, again, a big part of this portfolio and evaluating those I think we all agree have some methodologic challenges

and are significantly different than process measures.

So the next one, again, related. PCI. A lot of patients with AMIs or angina or other risk factors undergo PCI. We actually are going to be looking at eight measures for PCI today. Two of them are new. Maybe four of them are new actually. And plus the existing measures. So, today that's going to be our topic. We haven't looked at these measures in awhile so the existing measures, if you notice there are three measures for mortality, one for inpatient, two for 30-day all-cause. So we will have a conversation about related and competing measures around mortality for PCI later this afternoon.

Next, I don't want to overlook the fact that NQS has a large portfolio of measures for coronary artery bypass graft surgery, but they are not for you to evaluate. These belong in our surgery portfolio and that committee actually will be meeting in the end

of May to look at measures not only of CABG but other types of surgery.

But be aware that we do have a goodly number of measures in this topic area significantly related to cardiovascular disease and being aware that they exist is important for your understanding of NQF's portfolio for cardiovascular disease.

So we do have some measures for post-acute rehab phase. So there are some measures for discharge after a PCI. And oh, my mistake. If you notice underneath the red PCI post-procedural optimal outcome therapy is an adherence to antiplatelet therapy. That's also got the large 2379 measure. That's a new measure and I forgot to highlight it in red. So you've got a couple of new measures.

Tomorrow we'll be looking at two measures of referral to cardiac rehab for both inpatient and outpatients. So we have a lot of things happening.

Then -- we're not done yet. So we

do have a lot of secondary prevention measures particularly for patients who've been hospitalized at both the hospital level and at the clinician level. And you can see a goodly number of various types of measures for various types of medications to be prescribed after that acute event. So again, large portfolio of measures.

Okay. So that was coronary artery disease and AMI. But that's not the only topic area where it comes to heart disease.

So we do have measures around heart failure. And so using that same patient-focused episode of care framework staff has drafted sort of a heart failure-specific framework using the bubbles.

And we really would like your feedback on this because again this is one we've drafted internally and we're looking to your expertise to help refine it. But again, it helps organize the framework into sort of a patient approach and how to think about

measures for heart failure rather than just a list of measures.

So next, the measures again -- again, we start with a population at risk. And certainly the smoking, the weight management, controlling high blood pressure. The weight management and smoking are in our health and well-being portfolio. Hypertension control is for you to evaluate though not at this meeting. And of course we already saw a large number of measures for coronary artery disease that could lead to heart failure.

So, evaluation and ongoing management for heart failure. We do have measures including one new measure that we will evaluate tomorrow on symptom and activity assessment. But again you can see these are both facility-level measures as well as clinician-level measures that are used significantly in CMS's measurement programs.

Next, again, acute phase hospitalization measures for heart failure.

You're probably all quite familiar with them. Again, report on Hospital Compare. There's a population-level admission rate that is part of our population health portfolio that you should be aware of.

There are the hospital-level measures as well as clinician-level measures. There is a new measure here for post-discharge appointment for heart failure. We also have the readmission rate that is being evaluated in our readmissions project.

We do have the 30-day all-cause mortality rate in this portfolio as well as an inpatient heart failure mortality rate. So, a goodly number of measures for the acute phase and outcomes as well for heart failure.

But we haven't ignored all other types of heart failure. So there are measures around rhythm disorders such as EKGs for patients with syncope. We have a couple of measures for atrial fibrillation, several measures for ICD use. Those will be coming up

in future meetings with you. All part of this portfolio though we won't be discussing them at this meeting.

Then we do have a couple of measures for cardiac catheterization, particularly one for children, an adverse event outcome measure. And so we don't want to forget heart disease in children when it's appropriate in our measurement.

And then the next one is the couple of measures we have for hypertension. One is a controlling high blood pressure measure. Another is blood pressure screening for adolescents.

And then I do believe we've finally reached the end of them. And we have a cost and resource use measure which is being handled by our cost and resource use committee along with other measures of cost and resource use. And it's a relative resource use measure for people with cardiovascular conditions across the spectrum usually associated with

hospitalization.

So, as you can see this is really one of our largest, most diverse portfolios. It's challenging to get your arms around this number of measures. And so we really would appreciate your input in terms of the best way to organize these measures. If this works for you, great. If you've got suggestions for revisions and improvement, that's great too.

So, any comments you'd like to make on the portfolio at this point I've got a couple of more things to talk about before we wrap up.

Any thoughts from anybody on the portfolio? I know I kind of ran through it relatively quickly. But I guess just any thoughts as you're undertaking and taking on this challenge? Yes, Tom.

DR. JAMES: This may be more of a parking lot issue, but I know that -- and Helen can jump in on this one. That NQF is dealing with the social determinants of health

and how those impact these various measurements.

As I mentioned, I come from a Medicaid company and so for that reason it's something which is important to our population. I want to ensure that it's someplace within our view.

DR. BURSTIN: Yes, thanks, Tom. So NQF embarked about six months ago on a body of work looking at whether outcome measures in particular but not exclusively should be adjusted for sociodemographic determinants.

And ultimately the report that came out which is still in process, I want to caution everyone of that, indicated that for certain outcomes where there's a clear conceptual relationship between the outcome and the sociodemographic characteristics, and those factors are not directly related to quality of care, and thirdly, there's an empiric relationship as demonstrated in the analyses some of those perhaps should be

adjusted. It was actually quite a nuanced recommendation.

It also said clearly that for measures where you're really interested in disparities and quality improvement those measures should be stratified. That report is still in process and actually comment closed last week. We have 667 comments to review, an NQF record. I don't know if that's good or bad, but it clearly I think shows we've picked a question where there's been a lot of consternation for a lot of years. So, the committee will have an opportunity to review those comments. Whatever happens with that it will all play out sometime in June. So we'll come back to this question if we need to depending on where we are with the report.

MS. MITCHELL: So, I know that there's been some work on multiple comorbid conditions. But I'm curious about how cardiometabolics is being handled and what sort of is the purview or not of this

committee.

DR. WINKLER: I know that the conversation comes up a lot. I'm not aware that we have any specific measures. Nor -- and this would be a good help from you all if you know if there are any in development around sort of metabolic syndrome and that particular risk group that certainly would be something that I don't think we've got any measures on but it sounds, you know, it would certainly be an important area.

If you're aware of any measures in development we'd certainly want to hear about it. Because again, I think you're right, you bring up an important gap area.

DR. BURSTIN: And just to build on that comment, Reva. I think the other issue is there's still a fair amount of lipids for diabetes, lipids for hypertension as opposed to really a more holistic view of all the different patient populations that should be part of a measure.

And it's interesting, when Reva made the point and there's only one hypertension measure, that's actually somewhat by intent. Many of the other measures have fallen to the wayside as that one measure has become more of the de facto standard that's used by CMS and all the federal agencies now. It has all the different sort of characteristics and different patient populations built into that measure.

So a lot of what you'll be talking about over the next couple of days as well is does this really need to be this disease-specific measure to Kristi's point, or is it really more of a global measure we should really push the developers to ultimately move towards more of a population view of who should be getting what, when for the best possible outcomes.

DR. SPANGLER: Related to that, and maybe an extension of Kristi's question. If there is a cardiometabolic measure, I mean

would that be our purview? Would it be the endocrine steering committee? Are there measures that -- and I guess the broader question. Would there ever be measures that are addressed by two different steering committees at the same time?

DR. WINKLER: Yes, I mean we're struggling with the best way to deal with that. As I mentioned, sometimes the assignment to projects can be arbitrary and that's why we want you to be aware of the crossovers. So, actually, you know, we would take a look at actually how it's specified to see what would be the most appropriate place to put it.

But I think we need to really think a little bit more about measures that probably belong in two places. Because again, our topic areas are somewhat arbitrary. We're having to make some sort of cut points.

But perhaps there may be a way of getting input from both committees so that

there was sort of a shared accountability for the measure perhaps that we'll have to talk about how that might happen.

But again, we would want input from both certainly. And not, you know, keep ourselves in our little silos. We're trying to break them down actually as much as possible.

Any other thoughts from anybody else? Yes, Liz.

MS. DELONG: I just want to put on the table that I'm a little concerned about harmonization. As we grow the number of measures and they cross different venues to make sure that we're not coming up with measures that are somewhat inconsistent. And that's what I'm particularly concerned with. And I just want that to be on the table.

DR. WINKLER: Sure. And I think given the large number of measures in the topic area I think that it really is a good argument for the need for harmonization. If

we expect all of these patients to be subjected to measurement and different unaligned non-harmonized measures are being used it just creates a bit of chaos for absolutely everybody.

So, as we look at measures I think you're aware as we discussed them in our workgroups that the issues of related and competing measures and harmonization is something that really is our fifth major criteria. And we do particularly want to focus looking at that.

And really to the degree possible get measures harmonized to facilitate their implementation out there. So thank you, Liz, for bringing it up because indeed it is a significant priority for us.

DR. BURSTIN: And just one more point on Reva's point. It doesn't always have to just be harmonized. Sometimes it's okay to say it's been measured in one setting and it's kind of done there and it's time to move on.

It's topped out.

I think there's a lot of measurement burden out there. Many of you on the front line of health systems know this all too well. We really need to be measuring the right thing at the right time. And if it's past due and it should be in a different setting or work across settings just have it done once, the right place, that's really important too. I think we really want to as much as possible reduce the measurement burden out there.

DR. HOLLANDER: I'm not sure it's the purview of this committee and it doesn't look like it pertains to these measures, but I'm gathering from what you said there's an admissions/readmission group.

And I just want to put out there for thought that what people have done with all these readmission measures is simply gamed the system. And now everybody's going to observation. And it represents the exact same

failing as the health system as when they get hospitalized and go upstairs.

And I think it's important as measures deal with readmission or admission that it actually consider observation as part of the admission pathway rather than just an excuse to not be counted in the measure. And I think that's getting lost.

Some hospitals are now admitting 50 percent of their patients to observation. So it's a problem that just needs to be addressed, although it may not be any of the measures we're talking about today and tomorrow.

And it was a major part of the discussions last round as well as this upcoming round. I mean, at least the analyses CMS has done would suggest the rate of decline of readmissions is not due -- it's really to the change to a lot of people being admitted to obs. But obviously we've seen a lot of that shift in the marketplace.

DR. WINKLER: Again, we fully expect that during the course of your discussions you will bring up thoughts and ideas that will prompt you in terms of gaps or why not measure this instead of that. Feel free to please throw those out there. We'll capture them and include them as part of your sort of assessment of the overall portfolio.

Just to kind of finish things up. As I said, this is a particularly important area for the National Quality Strategy to reduce morbidity and mortality according to cardiovascular disease.

So the question is how well are we doing. And so I think one of the best sort of measurement tools to get a global population view which is an important sort of bellwether is the National Healthcare Quality Report.

And so I just picked the most recent report to kind of ask the question how are we doing. And so they report on three different measurements that I think are

particularly salient for us.

One is around blood pressure control, and as Helen mentioned blood pressure control is the measure we are looking at. And so I think that over time looking to see how we are doing as a nation by age group over the last decade we are getting a sense that things are improving and that's great.

But if you notice the highest is still only 50 percent. So we've got a long way to go.

And I think measurement and the measures we have are some of the tools to help us continue to improve. By no means is it the only thing that's driving improvement. It's actually the work that's being done on the front lines with clinicians and their patients. But nonetheless, it's useful to keep an eye on how are we doing globally.

So the next one, again, deaths from heart attack. Something that's improving significantly. This again per 1,000

admissions. Because the overall incidence of MI seems to be declining as well over time. Perhaps we're getting -- intervening up front into the risk factors before an AMI actually occurs. So there does seem to be improvement in mortality around heart attack. So we hope that this continues.

And keeping an eye on those outcome measures is really an important sense of how we're succeeding in improving the quality of care in this particular topic area.

And then lastly, hospitalizations for heart failure, again a chronic condition that tends to be progressive at either a greater or a slower rate. It's interesting that when it's stratified by age groups we're down low for everybody but the Medicare population which of course this is a huge area.

There does seem to be over the last decade some decline in admissions for heart failure. We certainly would like to see

ongoing improvement to decrease cost to both patients and to the system. So I think that in general we have a sense that improvement is occurring. And so keeping an eye on how we're doing over time will help provide the greater context for the portfolio.

And perhaps as we look at the measures to see -- get a better understanding of what might be the greatest drivers for improvement across the nation.

So that's the last one for me. And I think we're getting ready to talk about what we came to do. Sana, question.

DR. AL-KHATIB: Actually, I have two questions. The first question is part of the National Quality Strategy is to address disparities. And I wanted to ask you, the existing measures that you shared with us today with the portfolio, are people required to report on all these measures by age, gender and race? Or is that measure-specific?

DR. WINKLER: It tends to be

measure-specific and more importantly program-specific. Because whoever is implementing and using the measures ultimately makes the decision on how they're reported.

I do know that for many of the measures in this topic area we have been given data from the measure developers by various strata to see how performance is among the different subpopulations.

Whether that's actually translated into the implementation programs really is up to the folks that are implementing them. So, that tends to be something that tends to happen after sort of the NQF endorsement.

But the conversations around appropriate use is something that we tend to have here at NQF a lot. And we would certainly encourage it.

And certainly we want to look at measures that demonstrate significant disparities in performance. Some measures, not so much, but some measures really we do

find that there are significant disparities and we do want to identify them and highlight them as having that kind of disparity. And that perhaps the measure is useful to be stratified to identify them in however it's being used.

DR. BURSTIN: I'll just add that as part of our work on disparities over the last several years we actually came up with a protocol for identifying which measures are disparities-sensitive. And we'll work through that.

Once you have approved a set of measures we'll go back through, identify whether there is a quality gap, how large is it, the prevalence of the condition in different populations and bring that back to you for your consideration.

We did this work, staff reviewed about five or six hundred measures already, identified a set of them as being disparities-sensitive. And we can try to highlight as we

go through which ones have already been identified that way.

And the hope is the measure has been identified and disparities-sensitive. There's a known quality gap, high prevalence in a population, that they should routinely be stratified going forward.

Now, whether they're stratified and used that way in terms of the federal programs or payment, but at least being sure that they're being used that way for disparities reduction and quality improvement is critical. So thanks for that question.

DR. AL-KHATIB: My other question has to do with all the measures that you showed us in our portfolio right now.

If we as a group discuss a new measure, that we consider a new measure that we see a lot of overlap between the measure that we're discussing and the existing measure, but we really see more value in the measure that we are considering, could we make

a recommendation to implement this measure and perhaps retire the existing measure? Or how does that work?

DR. WINKLER: I think it will go along with your conversation in terms of the recommendation for this one.

The other measures will come up for your review. And so we're wanting -- we will want to take note of that so when it does come up for its next review. It's like guys, remember last time you maybe not so much on this one, you've done the other one.

So, again, because we get to stay together for a couple of years we have this opportunity that we didn't have previously. And we're really seeing that as some of the value of a standing committee who can actually work that way.

We would not want to automatically retire a measure until it's had a chance to go through its appropriate time to review. But certainly we'll want to carry forward any of

these conversations that you might have about similar related measures to say take note, next time we see this measure perhaps it won't be as important in light of this new measure.

All right. Thank you all again. All your feedback during the meeting, after the meeting. If you have a thought somewhere down the road feel free. We really want to work with you to help guide this portfolio to be as good as it can be and most useful.

And so particularly you folks out on the front lines who really are being measured and using measures can give us a lot of good insight on how effective and impactful these measures are going forward.

So that's it for me for right now. Wunmi, you want to kind of get us ready and we should be able to get started looking at measures.

MS. ISIJOLA: Sure. We're going to start off.

But just some of the ground rules

for today's meeting. We do expect you to have reviewed all of the measures, not necessarily just the measures that you are discussants for. And really basing your evaluations and recommendations based on the criteria at hand. And that information was shared with you and is also available on the SharePoint site.

Remaining engaged at all times. Obviously at any point in time you can excuse yourself to the restrooms. And making sure that your comments are concise and focused. We are going to turn it over to our co-chairs and they're really going to facilitate that.

But in terms of effectively and effective discussions we ask that you use your name tent cards and stand them up vertically. And our co-chairs will call on you accordingly.

We do have two of our committee members on the phone and they will be participating in the discussion. And Lindsey will make sure that that happens.

Cathy, our speaker, could you ensure that Jeff Burton's line is open?

OPERATOR: He has not dialed back in at the moment.

MS. ISIJOLA: Okay. He's not on yet. Okay. Well, in that case I will turn it over to Dr. Kottke and Dr. George and we will begin with our first measure.

DR. KOTTKE: I'll turn it over to Mary in a second here, but if I could draw your attention to a couple of things.

The first measure has been given an hour discussion which is twice as long as all the other measures.

And we have a lot of measures to do. As Reva said earlier this morning to me we could discuss them endlessly but we're not going to do that.

And you were very nice at keeping your bios real short. And I'll let you know that Mary and I have agreed that we're going to run on time.

And there's also voting. Within that 30 minutes there's four votes. And so please -- if there's something very important please say it. But if it's just a tangential espousing on something you want us to know there's beer for that. And so.

(Laughter)

DR. KOTTKE: Right, but there's no beer at the meeting and so it'll have to wait until after 6 o'clock. So I'll let Mary call for the first measure.

DR. GEORGE: Thank you. So our first measure is number 0964.

DR. WINKLER: Mary, let's bring our developers up to join us at the table.

MS. ISIJOLA: Okay. So really how it's going to run is we're going to have the developer come to the table and give a brief description of their measure, really brief.

And from there we'll ask that the lead discussant kind of speak to each criterion. And if there are questions we can

most certainly direct it to the co-chairs and they can facilitate that process.

But we really want to make sure that we're talking to each criteria as it was laid out. And we have provided you guys with the discussion strip so you can follow through that process as you present your measure.

And with that being said I will turn it back over.

DR. GEORGE: Okay, and the measure developer, Dr. Massoudi. Welcome.

DR. MASSOUDI: Good morning. Thank you, Drs. George and Kottke. And thanks for having me here. I'm sorry I'm leaving this evening. It sounds like the beer would be fun.

I'm Fred Massoudi from the University of Colorado. I'm a senior medical officer of the National Cardiovascular Data Registries or NCDR upon which this measure is based.

Again this is therapy -- this

measure is 0964 therapy with aspirin P2Y12 inhibitor and statin at discharge following PCI in eligible patients.

And I'm going to be very brief and would be happy to answer questions. But essentially this is a composite measure, a guideline-based medical therapy with three classes of medications following PCI. It's an all-or-nothing composite.

It includes the aspirin P2Y12 inhibitor, so the clopidogrel family, and statins. Each of these therapies is a 1A guideline recommendation in hospitalized patients through PCI.

It's an all-or-nothing composite. That is to say each patient has to be treated with all the medications for which they are candidates and is reported on the hospital level.

The feasibility and reliability and validity have been tested fairly widely in the CathPCI registry. It's been used as part

of feedback within the registry for the last three years. The registry is now used in 1,600 U.S. hospitals and has collected data on more than 14 million patients.

This measure again is fed back to sites as part of the executive summaries and will be part of the voluntary public reporting program sometime this year.

I would add that this is a renewal of a previously endorsed measure. This was discussed at the last panel. Actually we had submitted each of the three individual components of this measure and in response to NQF's requests we made an all-or-nothing composite of the three individual components into this all-or-nothing composite measure.

And so at this point, I don't know if that's what you were looking for, but.

DR. GEORGE: Thank you. And we'll go ahead with the primary discussant.

DR. AL-KHATIB: Okay, well, thank you. First, I do want to take a minute to

thank whoever sent this template of how we're supposed to lead the discussion because it really helped me organize my thoughts in this whole response.

So, as was stated, this measure has to do with looking at patients undergoing PCI looking for those patients who are receiving prescriptions for all medications, namely aspirin P2Y12 receptor inhibitor and statin at discharge following PCI. The level of the analysis is the facility or the hospital.

And as was stated this is a composite of three process measures. And this was a request by NQF as we were reminded of that on the phone and again today. So thank you, Fred.

In terms of looking at the evidence here this composite measure as I said has three process measure components in terms of providing the support and the evidence for that. They based it on guideline

recommendations. In fact, several guideline recommendations as well as a 2013 JAMA systematic review that included 91 publications with priority given to data from large randomized controlled trials, systematic reviews and meta-analyses. So based on these data I actually rank the level of evidence as high for this particular measure. And I'll open it up to others to chime in.

DR. GEORGE: Are there any additional comments from the secondary reviewer?

DR. CROUCH: My only comment was that there was a lack of empirical validation of the composite measure. It was an expert consensus view as opposed to database.

The validity data was good. The reliability was based on expert opinion.

MS. TIGHE: Sorry, we're just talking about the evidence criteria right now.

DR. GEORGE: If you follow along with the script that was sent we will be going

step by step and taking a vote after each. So, at this point we'll open it up for discussion on the evidence.

DR. WINKLER: I guess the question would be we have three components and we're going to have other composites so I just wanted to bring this up is when you have a composite measure of the components we really want to look at each individual component measure and the evidence for each of those. So I think Sana did address that, but that is an important aspect when you're looking at a composite measure.

MS. ISIJOLA: Are we ready to vote?

DR. JAMES: Just one question. I mean, certainly as a general internist I subscribe to this. You always have to keep in the back of your mind as somebody who writes scripts for a living on weekends that what is the interaction of the various meds.

Clearly I saw there was evidence

for the use of the platelet and platelet drugs, and there's clear evidence for the use of statins. But what is the evidence for all three together? Did I miss that in here?

DR. MASSOUDI: I think as with many of these things there's not a lot of great evidence for any of those things in any field for any therapy. And so I don't think -- there's not a randomized controlled trial that compares to one incrementally over the other two.

However, each of these components is based on a pretty widely accepted class 1A guideline recommendation.

DR. HOLLANDER: So, I guess since this is the first time we're discussing composite outcomes we know there's data on -- and this follows on Tom's comment I think. We know there's data on layering on the antiplatelet agents to aspirin. It's not really clear to me there's really good data on layering on statins to those two.

So I have no issue with this. And I understand from our prior telephone conversation that NQF asked for the component. So I get that.

But my question is do we want composites to be therapies that have been tested together, or is it okay for composites to be independent therapies all of which have guideline recommendations?

DR. BURSTIN: I'm happy to respond especially because we encouraged this the last time. They are independent therapies that all individually have evidence.

And I think the key thing for structuring it and we thank ACC for doing this as an all-or-none composite was the idea that simply doing -- adding each one on incrementally was not enough. You actually want to in fact see there was evidence that in fact doing all three was really important.

But I don't know that we need to have evidence of the specific additive. I

think it's really just that all three of them are critically important in terms of evidence.

DR. KOTTKE: And if I could comment. One reason for the composite is that you could look at each one and have a score of 85 percent on each one. But when you look at it the perfect care score is really pretty low. And that's one reason that we developed the composite at HealthPartners is to drive the bar upwards in terms of quality of care.

DR. PHILIPPIDES: Quick question. Does it mention ticagrelor? Or is it just --

DR. AL-KHATIB: It does. The initial document that was circulated mentioned ticlopidine and then we clarified that that was a typo. They meant to include -- yes, exactly. So it has been revised.

DR. WINKLER: So now we get to vote. Go back one, will you? Go back to the beginning of the voting slides. Okay.

We just want to give you a sense of sort of how voting goes. And for those of

you who have worked at this before voting has changed a little bit around what is the aim to each consensus.

And so in the past it was -- 50 percent plus one was enough. But we've had a lot of feedback that said you know, that really isn't -- that's kind of iffy.

So we've changed the voting results, evaluating the results in the following way. Anything above 60 percent of the committee passes. That's what it takes to pass. If it's less than 40 percent, it fails.

But the 40/60 corridor is an area where it feels the committee really hasn't reached consensus. And so there's a consensus not reached.

And so realize that that puts us in a bit of a holding pattern in terms of not pass or fail. And we'll continue evaluating the measure to see if we can figure out where the consensus lies among the group.

Otherwise, if a measure fails on

certainly any of the importance criteria or the scientific acceptability criteria we just kind of stop because those are must-pass.

If it's in this sort of consensus not reached corridor we will continue evaluating the measure till we get a sense of what's going on. Okay?

So this first measure as we go through, it's the reason we gave you an hour so that we could talk through all these nuances around voting and what the various votes mean. Okay? So that's how we're going to count the votes.

Also, a quorum is important. We need 75 percent of you. So that's why the staying with us except for breaks is really important, so we don't lose people. And we realize tomorrow afternoon as you go to catch your flights we've got time pressure. So we'll be paying attention to that as well.

So, when we're looking at evidence it is part of the importance to measure and

report criteria. And evidence is the first but we'll talk about performance gap and priority.

And then in a composite measure we're going to look at the construct of the composite. So those are the subcriteria for a composite measure.

So we were just talking about evidence. All right, let's go to the next one. It's not an outcome measure, it's a process measure, so we go to the next one.

Okay. It seems complicated. I hope it's not terribly complicated. I hope you've had a chance to look at the algorithms.

And so, based on your review of those algorithms you have five voting options. Sometimes too many choices is difficult.

You can rate the measure high on evidence if indeed you have the results of the quality, quantity and consistency of the body of evidence. In other words, it's golden. Okay?

Moderate is still a passing grade but perhaps you don't have the details of quantity, quality and consistency, or perhaps if you did and the evidence isn't as stellar as you might like.

Low means you have information but the evidence really does not support the relationship to outcomes.

And that is distinct from insufficient evidence where you don't have a lot of information around it. There just isn't evidence to deal with.

So if there's insufficient information you actually have two choices. If it's insufficient and you're comfortable saying measure goes down. There just isn't enough evidence to support it and we don't want to go any farther.

On the other hand there are rare but occasional measures where there isn't strong evidence or much evidence at all, but yet the committee feels that despite the lack

of evidence it's still an important measure and they feel comfortable holding people accountable for something with no evidence. So that's your option number 4, insufficient with an exception.

So, if you look at the algorithm you'll see that there -- you could be led to any one of those options.

So, before we actually ask you to vote you should have a copy of the algorithm in front of you. I'd like you just to take a look at it. And does anybody have any questions on how that algorithm works? Because we were asking you to use that to refer when you're doing your voting.

MR. BURTON: This is Jeff Burton. Can you hear me?

DR. WINKLER: Yes, hi Jeff.

MR. BURTON: Great, great. I do have one question. When I was going through the algorithm there were points for some of the measures where I came to the very end and

based on the criteria for QQC I was kind of at one point. But then if I look at another statement here it says that if you feel that there's moderate certainty that there's a net benefit outweighs the harm.

Could it be either/or? So if it's something that has a low consistency in the systematic review and you'd actually rate it as low according to the algorithm.

However, the overall body of evidence and common sense would maybe lean towards a moderate vote because there's more certainty that there actually is a decent benefit there. How does that work?

DR. WINKLER: Again, if this were strictly sort of a one-two-three calculation we wouldn't need you all. So, we're asking you to help us find the best answer here.

So, if indeed your systematic review, really the conclusions are there's no relationship, or the -- there's too much uncertainty, we really can't support that.

Then you're going to rate it accordingly, moderate, or low as the evidence just isn't there.

In your hypothetical, Jeff, I'd have a hard time understanding a situation where a good systematic review came to the conclusion there was no evidence or low everything to support the relationship, and yet you feel there would be a moderate level. Somehow that's a little inconsistent. We need to talk about it further to get a better understanding of what exactly those discrepancies are. So that's the best I can help you with right now until we have a real example to talk about it.

MR. BURTON: Sure, sure.

DR. WINKLER: All right. Does anybody have any questions about the criteria and about the ratings for voting? Yes, Liz.

MS. DELONG: We're voting on the entire composite for evidence?

DR. WINKLER: Correct. For

composite you're talking about the entire measure. So, again, it will be the specifics of the evidence for each component in aggregate. Because you're talking about the entire composite measure.

MS. LUONG: So, by now everyone should have received a voting fob. Please let me know if you have not.

And how we're going to do this is I will start the timer. You'll see the timer on the right corner of the screen. And if everyone could just point their fob to me and click the number of your choice later when I start it that should be it. And we will start right now for evidence 1A. The timer has started.

Can you try pressing it again? No, it doesn't double-count. Technical difficulties. Excuse us real quick.

We have 13 for high, 7 for moderate and 1 for low.

DR. AL-KHATIB: Moving onto

opportunity for improvement. The developers shared information and data with us about the gap in care. They shared an 88.6 percent use of all three medications in patients undergoing PCI.

And at least during our phone call one of the participants didn't think that that was a big gap in care. And I actually agree that it's not a tremendous gap in care.

But I think we should try to shoot for close to 100 percent in these patients. Because again, when we talk about the specifications we will be excluding people with contraindications. But because we want this draft to be as close to 100 percent as possible I considered that a gap in care significance.

MS. LUONG: So we will continue.

DR. KOTTKE: Secondary. Any other discussion? Fred, how did you a sort of frustrating portion of people just don't tolerate statins.

DR. MASSOUDI: So, great question. So if there was a documentation of a contraindication for -- let's say a patient comes in and they have a contraindication for a statin but not to the two antiplatelet therapies. They would pass the measure if they receive the two antiplatelet therapies and the statin wouldn't be considered because they had a contraindication for that therapy.

DR. KOTTKE: So a patient report of --

DR. MASSOUDI: So, a documented contraindication along the lines is what's done with the CMS medication measures.

The other thing I would just for one moment just clarify in terms of the distribution as well. The lowest 25th percentile was an 83 percent. The lowest 10th percentile in a 76 performance rate just for perspective.

MS. DELONG: I have a question. Was this generated from the NCDR? How

representative is this 86 percent?

DR. MASSOUDI: So it's generated from the CathPCI registry. As we've discussed before the CathPCI registry is used in about 1,600 hospitals which is well north of 80 percent of hospitals that do PCIs. And again has been reported in now 14 million patients. Not this specific measure but since the onset of the registry more than 14 million patients. At this point probably represents around 90 percent plus of patients undergoing PCI in the United States.

MS. MITCHELL: To follow up on your -- I think you presented a range. What's the range of performance on this measure?

DR. MASSOUDI: The range from the 1st to the 90th is 55 percent to 96 percent.

DR. CHO: So, let's say there are other measures like aspirin and P2Y12 inhibitors and statins, all separate measures. So if we vote for this do those measures go away?

DR. MASSOUDI: Those measures actually don't exist as an endorsed NQF measure. Again, we applied as individual measures the last time. We're instructed to include it into all-or-nothing composites. So those have not been endorsed and we're not asking for endorsement for the individual component measures at this time.

DR. AL-KHATIB: I just wanted to add one comment about disparities because I was expected to cover that as well under opportunity for improvement.

And one thing that the developer stated in the submission is of particular interest is that when compared with the expected mortality rates those with private insurance had significantly better survival, while those with all other insurance types did worse. And then they talked about some geographic variations as well.

I wanted to ask if we ever thought of using insurance status for reporting.

Could we require that data be reported based on insurance status?

DR. BURSTIN: It's actually fairly common. For example, if you see some NCQA measures they routinely report by commercial, Medicaid, Medicare. So I don't know if the differences are there. It's certainly something you could talk about with ACC.

DR. WINKLER: Fred, you mentioned that this measure may become part of a voluntary reporting program for ACC. What are your thoughts on reporting and addressing some of the disparities questions?

DR. MASSOUDI: At this point the measures are reported. They're not specifically stratified by various socioeconomic status or insurance status. It could certainly be performed. And the details of how this will be presented and reported are still in development as you can imagine, as we've discussed on previous calls.

And just, I should have introduced

Lara Slattery from ACC staff before. I apologize, I got so excited about presenting the measure that I forgot to. So, please forgive me. Thanks.

MR. BURTON: Was there any evidence -- I don't think I saw any target performance level that could be set for this measure. As opposed to comparing against the mean. I know that mean for all the hospitals was in there. But was there any target performance level that was established through the evidence?

DR. MASSOUDI: So, as Dr. Al-Khatib pointed out this is an all-or-nothing composite based on three medications. For each individual patient they would be excluded from that particular medication if they had a contraindication. So the ideal target performance level on this measure as it is for many process measures where you exclude patients with contraindications is 100 percent.

MR. BURTON: Okay.

DR. GEORGE: Are we ready to vote? Any other discussion?

DR. WINKLER: For this your voting options again are really just your qualitative assessment of high, moderate, or low. High or moderate will be a passing grade. Low will not. If you feel there isn't any evidence or insufficient data for you to make a determination that option is available.

So you will see this generic high/moderate/low voting scale on several of the criteria.

MS. LUONG: The timer has started. We have -- for criteria 1B 8 for high, 13 for moderate and 1 for low.

MR. BURTON: I just want to confirm that you're getting my vote over the chat, the webinar chat. This is Jeff Burton.

MS. LUONG: I am. I am. Thank you.

MR. BURTON: Great, thanks.

MS. ISIJOLA: So we'll move onto priority.

DR. AL-KHATIB: So in terms of priority I absolutely believe that this measure addresses a significant health problem. CAD is a very prevalent condition. PCI is a very commonly performed procedure and is associated with high costs. And we really need to ensure that we optimize the use of evidence-based therapies that have been shown to improve survival, reduce risk of infarction, what have you. So multiple outcomes that could be improved with the use of these therapies. So, for me I think priority is definitely there.

DR. GEORGE: Any discussion on priority? We're ready to vote on priority.

MS. LUONG: The timer starts now. For criteria 1C we have 18 for high and 4 for moderate.

DR. WINKLER: We have one more -- because it's a composite we need to go to the

composite construct in terms of this 1D. You'll only see this on composite measures. It's the final criteria under importance and it's the quality construct of the components, the rationale for putting them together. And then any aggregation or weighting issues. So it's all about does this composite construct make sense.

DR. AL-KHATIB: So for the construct I ACC would argue that it's pretty good, pretty reasonable and logical. We all have concerns about component endpoints and measures because we always raise the question as to what kind of weighting system you're using.

And I don't know that you can ACC justify that or defend that in association with any composite measure.

But with that caveat in mind I actually think that the construct is pretty good.

DR. GEORGE: Any discussion on the

measure construct? All right, we'll vote on the measure construct.

MS. LUONG: The timer starts now. All right. For criteria 1C we have 18 for high and 4 for moderate.

DR. GEORGE: So we'll move along to the scientific acceptability.

DR. AL-KHATIB: Okay, so in terms of the scientific acceptability, just a summary of the specifications. The numerator is all patients undergoing PCI who are eligible for all these medications, aspirin, clopidogrel, prasugrel, and ticagrelor and are prescribed those medications.

Denominator is all patients undergoing PCI who are eligible for all of these medications, meaning they don't have any contraindication to any of those medicines. And I forgot to mention statins as well.

Exclusions are death or presence of a contraindication. And the measure uses the CathPCI registry. This was described

briefly by Fred.

I personally don't have any concerns regarding the specifications, definitions, or coding.

DR. GEORGE: Any discussions on the scientific acceptability?

DR. CROUCH: There's a harmonization issue with this one and the one we're going to discuss next with the exceptions. I don't know which way we want to go, whether we want to leave that till the next one or bring it up now?

DR. WINKLER: We'll talk about that after we've talked about both of them. And then we'll talk about that part.

DR. CROUCH: Okay.

DR. WINKLER: But thank you for bringing it up.

MS. DELONG: I have a question. I'm a little confused. I thought Fred said that if somebody was contraindicated to one of the measures, one of the components but not

the other two that person would still be in the denominator and the numerator would ignore the contraindication.

DR. MASSOUDI: Yes, that's correct. So a patient who's eligible for any one of the therapies would end up in the denominator. And if they receive treatment for all the medications for which they were eligible they would count in the numerator.

MS. DELONG: Okay. So that's different from excluded if they're contraindicated to any of the measures.

DR. MASSOUDI: Yes, they're excluded at the medication level if that makes sense. It would be excluded entirely if they had contraindications to all three medications.

MS. DELONG: All of them, right.

DR. MASSOUDI: They would be included entirely if they were. Yes.

DR. PHILIPPIDES: How were you deemed to be excluded because of inability to

take a statin? What was the definition of statin intolerant?

DR. MASSOUDI: Just as what's done with the, say for instance, the CMS Hospital Compare measures. The clinician documentation of a contraindication is considered a contraindication of that medication. Sort of a standard contraindication exclusion.

DR. WINKLER: Question just on that. Contraindications are captured in the registry?

DR. MASSOUDI: Yes, that's correct.

DR. WINKLER: Okay. Just one other question that got brought up in the workgroup was the issue around the age indication for patients for statins. The workgroup brought it up. About age 75. No? Okay, not a problem.

DR. MASSOUDI: Just to clarify. We document that a contraindication was present. We don't catalogue the actual

contraindications, but just whether or not a contraindication was present.

DR. JAMES: Just to reiterate the discussion that Judd and I had whether there is scientific evidence of the combination of all three. I think it's very clear about each individual component. It's what happens in the individual person to handle three different medications.

DR. MASSOUDI: You know, again, I think it's an issue with any composite measure. I would say that there are probably numerous examples of composite measures where each of the individual components are evidence-based. But there's not evidence for additive benefits with specific agents.

Having said that, however, I would say that the more contemporary secondary prevention statin trials are trials of statins over and above antiplatelet therapy standards currently.

Again, also the guidelines clearly

recommend as class 1A recommendations the use of these three medications in conjunction with one another in patients who might have done PCI. So this is all guideline-based on a class 1A recommendations for the use of all these medications together.

DR. GEORGE: Any other discussion? If not we'll go to a vote on the scientific acceptability.

DR. WINKLER: Did we talk about the reliability testing which is part of reliability?

DR. AL-KHATIB: I don't think there's a vote right now. We keep going. So, reliability.

So in terms of the reliability testing what the developer did is empiric testing using the CathPCI registry with data from 1,386 hospitals.

And testing was done at the data elements level, not the measure score level. And then they talked about reliability testing

was performed using correlation of random split halves of the participating hospitals. And talked about the correlation between the two being pretty high at 0.92.

They also highlight all the quality improvement and assurance within the NCDR that includes onsite audits and interrater reliability assessment conducted to validate the audits. I actually have seen those data although I don't know that the results were included in the submission but certainly the data are very reassuring.

So in terms of reliability testing I believe that what they showed demonstrates that the measure data elements are repeatable, producing the same results. A high proportion of the time when assessed in the same population in the same time period. So I think overall it's pretty good.

Based on the algorithm that was shared with us if they did not do the testing at the level of the measure score which I

actually didn't see that. I'm sorry if I missed it. I only saw the one that was done at the data element level. Then the highest ranking that this measure would get is a moderate ranking if they don't have a measure score.

Again, Fred, please let me know if I missed that. I do -- I did notice the data element testing but I didn't see one at the level of the measure score.

DR. MASSOUDI: That's not in the submission.

DR. GEORGE: Any questions on the reliability or further scientific acceptability?

MS. DELONG: This will probably pertain to a lot of these things when we look at reliability. The correlation says something but not everything. It would be good to see the percent agreement in the -- on the diagonal cells. Because then you have an idea of how many patients they said yes in one

group but no in the other. I think that's important information and I'm not sure we're consistently getting that in the reliability.

DR. MASSOUDI: I'll bring your attention to -- there are a couple of figures. Again, this is on the composite, so not on the individual components, again. Because we haven't focused on those because we're not applying for endorsement for any of those measures.

But the figures show the first and second sample validations as you can see, you can see the correlation, the composite there in Figure 2 between those first and second samples.

I don't know if it helps but in my document which I think would be similar to yours it's Section 2A2.3.

DR. KOTTKE: Does anybody have any overwhelming concerns about this? I'm looking at Figure 2 in my document and it is blank.

DR. MASSOUDI: Funny, it's on page

56 I'm on, but it's in the mid-fifties.

MS. DELONG: I'm sorry if I've caused a lot of confusion. All I was looking for was a 2 by 2 table that had in this sample yes/no and the other sample yes/no. Are those dots sites then? What are the dots?

DR. MASSOUDI: Yes, those dots are sites.

DR. WINKLER: Just to clarify on this, are those results at sites for the measure result? Or the data elements?

DR. MASSOUDI: The composite measure.

DR. WINKLER: Result.

DR. MASSOUDI: Yes.

DR. WINKLER: So that's a performance measure score. So you do have testing, reliability assessment at the level of the measure score. Well, that changes the eligibility on the rating.

DR. GEORGE: Just to recap that. Having both data element testing as well as at

the measure level does make this eligible for a high rating.

Other discussions on scientific acceptability or reliability?

MS. LUONG: The timer is starting now.

DR. WINKLER: This is voting on reliability. We'll do validity next.

MS. LUONG: For criteria 2A 16 voted for high, 6 voted for moderate.

DR. GEORGE: Okay, we'll move onto -- you've got validity. Just wanted to make sure.

DR. AL-KHATIB: Okay, so there was no empiric testing of validity for this measure that I could find. What the developer mentioned was face validity was described as content validity of this process was achieved by the specialized expertise of various ACC committee members involved in the development or approval of the measure.

And they went onto say that we

believe the content validity of this measure has been achieved by virtue of the noted expertise as I mentioned. The individual components of the composite have already been shown to impact clinical outcomes. The empiric analysis demonstrating the individual component measures fit the overall quality construct.

Testing will focus on construct validation which will test the hypothesis on the theory of the construct that following these processes for patients undergoing PCI would lead to better outcomes.

This research is expected to ultimately be published in the medical literature. While the analysis will likely not be ready prior to the submission deadline of the cardiovascular endorsement maintenance project they will be available prior to the close of the measure cycle.

And that the analysis in preparation for publication can be provided

upon request or at publication. But that was the extent of what they mentioned regarding validity testing.

MS. TIGHE: And just to jump in, this would also be the point in time to discuss the validity of the specifications, whether they're consistent with the evidence. I believe we've touched on it but just if there's anything to raise at that point too.

DR. GEORGE: So any comments on threats to validity?

DR. AL-KHATIB: From my perspective although they did not do the testing that we would like to see I don't see any like major concerns about why the data or the process wouldn't be valid.

It would have been nice to have the testing to prove that, but knowing the CathPCI, knowing the process, what they're proposing here, I personally don't see any major threats to validity.

DR. GEORGE: Any discussion on the

validity? Comments on the phone? We'll move to a vote on the validity.

DR. PHILIPPIDES: Quick question. So basically this algorithm, am I correct that if they're relying on face validity and there's not been empiric validity testing then the highest level that can be achieved for this would be moderate? Is that correct?

DR. GEORGE: Yes. Okay, we'll move to a vote on the validity.

MS. LUONG: Timer starts now. For criteria 2B 2 voted for high, 17 voted for moderate and 3 voted for low.

DR. GEORGE: All right, so we'll move onto feasibility.

DR. WINKLER: There's one other criteria for the composite. And again, this is looking at the empiric analyses of the various composite aspects. Again, this is section 2D on the submission. Whether the components fit the quality construct.

Typically an analysis might be the

frequencies of performance of each of the subcomponents or any issues around aggregation and weighting from a testing perspective. So this is kind of the scientific acceptability of the composite construct if you will.

DR. GEORGE: Any discussion?

DR. AL-KHATIB: So the questions that I see here under 2D on the form is do the component measures fit the quality construct. And I would say yes.

Are the objectives of parsimony and simplicity achieved was supporting the quality construct I would say yes as well.

DR. GEORGE: Any other comments?

DR. PHILIPPIDES: Let's say for the sake of argument that everyone across the country gave the antiplatelet agents 100 percent of the time. But all of the play, the real difference in performance was in just one. Would there still be a good reason to pursue a composite measure?

Because my suspicion, and I can't

tell because I don't have data, is there's probably more wah-wah in the statin than in the aspirin and the clopidogrel.

And so would it be simpler and allow people to spend less resource and get to the same sort of benefit if we only focus on statin in this case? I'm just throwing it out there.

DR. AL-KHATIB: Well, the data that they showed from their study, you know, the testing that they did with CathPCI showed ACC variation, significant variation in relation to the use of all three medications. Probably much less so for aspirin but certainly they saw some evidence of variation for the P2Y12 receptor antagonists and statins. Less so for aspirin.

DR. WINKLER: George, the answer to your question is specifically the purpose of 2D is to answer exactly that around the quality construct. Because you're right, there could be a composite measure that's

driven solely by one component.

DR. GEORGE: We'll move to a vote.

MS. LUONG: The timer starts now.

MS. TIGHE: For Ted and Jeff we're voting on the composite 2D criterion, what are the component measures to the quality construct.

MR. BURTON: Yes, just submitted mine.

MS. LUONG: So for this criteria 9 voted high, 12 voted moderate.

DR. AL-KHATIB: Okay, moving onto feasibility. The data source as was stated is the CathPCI registry. And we raised this question during the call and Fred answered the question during the call.

And again he reminded us actually participation in the CathPCI is excellent with an estimate of about 90 percent of PCI that are taking place in the United States are being captured by the CathPCI. As such I ACC have no feasibility concerns.

DR. KOTTKE: Fred, can I ask a question? What's the demographic or the epidemiology of non-participation? Do we know? Is it sort of rogue, or is it organizations that are really stretched financially and operationally?

DR. MASSOUDI: So, I'll defer to Lara for a little clarification. I mean, in general it's hard to know what you don't know in a sense.

We do know they tend to be smaller sites. But Lara, if you have any elaboration on that I'd welcome that.

MS. SLATTERY: Sure. So, it does tend to be the smaller-volume sites or sites where there may be a state reporting mandate that differs from allowing to be able to participate in our registry and no other driver or funding within the facility to support them doing both types of reporting.

DR. HOLLANDER: So what would be the ramifications of this measure passing in

terms of cost to the hospitals that don't participate or consequences if they continue to not participate? And is there any insights as to whether those are underperforming hospitals or the same as every place else?

MS. SLATTERY: Well, we operate the registries as voluntary programs. And as we've mentioned part of the reason for seeking NQF endorsement of this measure is that it will roll out into a voluntary public reporting opportunity. So, while it is not our intent to disadvantage those sites the structure of our reporting out of our registries does mean those hospitals that aren't participating in our registry are not eligible for our public reporting voluntary option.

Beyond that we do not know anything about those hospitals unless they happen to be participating in a state that has a similar type of public reporting component to it. We do not personally track that

though.

DR. GEORGE: Other comments on feasibility? Tom?

DR. JAMES: Just a quick comment and that has to do with the phenomena of code creep, or measure creep. We often find when things get to the MAP that those being held accountable may go beyond those originally intended.

This is clearly a facility-based type of measure. It would be inappropriate for this to go onto a physician health plan or community as level of accountability. So I want to make sure that that's clear when it goes through.

DR. WINKLER: Tom, just to clarify, the specifications of this measure though are at the facility level. The next one coming up actually is the same measure at clinician level. I don't believe we have this measure at a health plan level however.

DR. GEORGE: Any other comments?

We'll move to a vote on feasibility.

MS. LUONG: The timer starts now. For this criteria 18 voted high, 4 voted moderate.

DR. GEORGE: Move onto usability.

DR. AL-KHATIB: Several things to touch on there. The measure is currently being used in a program called the Blue Distinction Centers for Cardiac Care. The sponsor is Blue Cross Blue Shield. It's not publicly reported, this is just a quality improvement with benchmarking.

The product brought to our attention that in July of last year they kicked off a program to give hospitals the opportunity to voluntarily publicly report measures. And this was not incorporated at that point but I think their plan is to include this measure in that program that they're working on.

In terms of information on improving performance over time they showed

trends where they found that there is proof of improved performance with the use of this measure. And of course the improvement in performance was significantly lower for the top performing sites. Certainly there was significant improvement in performance for the sites that did not initially perform as well.

And in terms of unintended consequences the developer mentioned inaccuracies of data collection and over-coding of exclusions. Certainly possible but I didn't see any major unintended consequences. So overall I felt that usability is pretty good.

DR. GEORGE: Discussion on usability? If not we'll move to a vote on usability.

MS. LUONG: The timer starts now. For this criteria 19 voted high and 3 voted moderate.

DR. GEORGE: And so I think at this point we move onto a discussion on

whether to recommend the measure for endorsement. Any further discussion? If not we'll go ahead and vote on overall suitability.

MS. LUONG: The timer starts now. We have 100 percent. Twenty-two voted yes.

MS. ISIJOLA: Well, I think we will break for lunch at this time. Thank you and we'll convene in about 30 minutes.

(Whereupon, the foregoing matter went off the record at 1:01 p.m. and went back on the record at 1:30 p.m.)

DR. KOTTKE: So despite a markedly different title this is a measure that's very similar to our prior measure. We'll let the ACC explain it.

DR. NALLAMOTHU: Hi, good afternoon. My name is Brahmajee Nallamothu. I'm a cardiologist at the University of Michigan.

The reason I'm here is I was the co-chair on the PCI performance measures group

that was sponsored by the AMA's PCPI as well as the ACC. And this is a measure that directly relates to the work of that group.

With me is Jensen Chu from the ACC. Any of the hard questions we will definitely kick over to him.

The nice thing about this measure is it follows on the measure that was just discussed in quite some detail. The difference between 2452 and 0964 can be really summarized by 2452 is focused on the individual level, the clinician level. And that's one point that I want to make up front.

So initially we're talking about a clinical-level measure. Again, thinking about composite medication use following PCI at hospital discharge.

The second thing that's an important part of that, and then I'll kind of stop and let the measure be discussed, is the key concept about harmonization.

The issue about harmonization is

there was some concern about a call held a few weeks ago in discussion of these two measures. I just wanted to kind of emphasize to you that that was a little bit of a mis-sight on our part. Both those measures conceptually as well as technically we see as being completely harmonized. And I can go into details as the discussion unfolds.

The last thing I'm going to just say is that obviously with this being a clinician-level measure I'm going to try to be a little preemptive in some of the discussion. I think the biggest issue is about attribution.

Obviously the way that we see it, and just to emphasize, we see that the last clinician who has performed a PCI, the operator is responsible for this measure.

There's all sorts of issues about this. And I'd be kind of interested to hear the discussion that happens today. But we do feel that this individual is very responsible

for both the initial prescription of this measure as well as its subsequent use of these medications in this population.

So with that I'll stop and interested to hear your guys' thoughts.

DR. KOTTKE: Okay, thank you. Primary reviewer?

DR. CROUCH: Just to reiterate it's the same thing as the previous measure except it's on the individual provider level attributed to the person who performs the PCI.

So as far as the evidence is concerned it's the composite of three things. The same issues that we've discussed before, same qualifications. I don't really have anything to add.

DR. KOTTKE: Okay, having nothing to add does anybody else have anything, any discussion?

DR. WINKLER: For consistency do you want to just stipulate your vote on the last one for evidence? Rather than re-vote.

DR. KOTTKE: Does anybody object to that?

MR. BURTON: No objection.

DR. KOTTKE: Seeing no objection we'll stipulate our vote on the last.

DR. WINKLER: We'll just carry the votes from the last time forward.

MS. DELONG: Excuse me. A couple of us are having trouble getting into the site. When I clicked on the measure it took me all the way out and I can't get back in.

MS. ISIJOLA: We are having trouble getting access to the Sharepoint site but we are working it internally to get it up and running again. So bear with us.

MS. TIGHE: If there's something you need us to send let us know though. If there's a document you're looking to reference during this discussion.

DR. KOTTKE: Opportunity for improvement.

DR. CROUCH: Opportunity for

improvement. The 25th percentile was 84 percent. The mean 88.7, the median 90.3. So there's modest room at best for improvement.

The bottom fourth have more room for improvement. The top three-fourths don't have very much practical room for improvement.

DR. KOTTKE: Any discussion. Does anybody feel they need to change their vote from the prior measure? Hearing none we'll just record the vote -- oh.

MS. DELONG: Sorry, I'm not maybe changing my vote but what is the variability in samples? I mean, some physicians treat very, very few PCIs, right? Do very, very few PCIs.

DR. NALLAMOTHU: I can speak to this briefly. It does come at some issues that I'm sure are going to be raised when it comes with the measure itself.

But in this sample we had about -- I think there was a little over 11,000, about 11,500 or so individual operators.

When we tried to do some reliability testing we obviously tried to include only those with at least 50 or more PCIs and that's currently the standard by which the ACC and AHA have considered volume requirements.

And in that group that brought down the group from about 11,699 to 4,064. But I do want to say a couple of things about that.

So one is obviously it suggests that there are low-volume operators.

The second is that -- I'm sure this is going to be in detail, but there are some concerns about the capturing of individual operator IDs within the NCDR registry which was the registry we rely on for some of the testing.

DR. AL-KHATIB: I just wanted to add to that because that was a concern that I had with regard to the testing that was done. And I know we'll get to that.

You mentioned identifying the actual physician. And in fact in the testing phase there was a great degree of missingness in relation to the MPI number. And so that's something maybe we'll get to in terms of what you're plans are to try to address this degree of missingness. But I think you bring up an excellent point.

DR. KOTTKE: Other discussion. So anybody need to change their vote? Hearing no comment, well, just should I ask for the vote? Okay, go ahead, Michael.

DR. CROUCH: As to priority it's the same issues as the other one. I don't see any reason for changing my vote. Anyone else.

DR. KOTTKE: Any questions or any further comments? Hearing none I propose that we just transfer our votes. Okay. Go ahead, Michael.

DR. CROUCH: Scientific acceptability --

MS. TIGHE: Sorry, we've got that

1D composite criterion. Whether the construct -- essentially whether the quality construct including the components make sense which mirrors the last discussion.

DR. CROUCH: Oh, sorry. Same issues as before.

DR. PHILIPPIDES: Right. So I guess I'll bring up the same issue. Does this, as a composite does it include all of the things that we feel that we should have in there for adequate post-MI care. And if there are things that aren't in it should we discuss why they're not in it?

And in regards to the elements that are there do we know how often they've been hit in general? Is there data to show how often people have done well with that measure?

So I guess I'm questioning as to how this composite was made and should we consider having a different makeup of it. Because this is, right, total optical medical

therapy for PCI, right?

DR. NALLAMOTHU: I can briefly comment on that. Again, the structure is very similar to the idea of the last measure. Picking three class 1A guidelines recommended therapies none of which have been studied kind of in unison and that idea of a synergistic effect but each individually. So I think the same issues that were raised before which I heard even over there and are all good points still hold in this situation too.

DR. PHILIPPIDES: Would we as a committee be better served in basically creating sort of, one, metric that basically takes into account all of the post-PCI care that a good facility should be doing, rather than have three on this metric and three on that metric. Should we take this opportunity to sort of bring that all together?

DR. KOTTKE: So for as an e.g. put cardiac rehab in there.

DR. WINKLER: George, just to

clarify, are you questioning the fact that there are only three components for this? And then you had another sort of question was do we need two measures, one at the facility and one at the clinician level that are different.

DR. KOTTKE: Leslie and then Sana.

DR. CHO: I agree with George. I think part of my concern is that there is a fair amount of measures and you have a lot of measures. And there's a lot of measure overload I think.

And you have, you know, the first initial measure, I understand that's a hospital base. This is a clinician base. But at some point you have too many measures that hits at the same thing.

DR. AL-KHATIB: I think part of George's question has to do with do we add anything to the measure. Like you know, maybe beta blockers, ACE inhibitors, what have you. And my understanding is that those are very well captured by other measures. And that --

so I assume, and please correct me if I'm wrong, that you did not include the beta blockers, ACE inhibitors, what have you, in this measure because you felt that those were very well covered in other measures.

DR. NALLAMOTHU: Well, I mean I think that there's two points to that.

First of all, I'm hearing a lot here. I'm not sure how much -- these are all great points and probably things that this committee needs to address at some point in time.

But the two things that I could say are the addition of other drugs. I do think that if you start to look at beta blockers after re-vascularization, particularly like uncomplicated single-vessel disease, it's going to run into an evidence base that's much more controversial. Same is true for ACE inhibitors, ARBs.

Again, this is not a total AMI population. If a person has an AMI and a PCI

they probably would be grouped in both groups. So to keep that separate. These are the three that we think are the most critical in this situation, had the most broadest appeal.

I don't know how to answer the question about cardiac rehab and all these other measures. We as a group just recently came up with 11 of them. I don't know whether you combine them all and create a super measure or not. I think that brings its own complexity to it.

The nice thing about sometimes teasing things out, and this is a real gestalt feeling, and you guys are going to be the ones that decide this, but the nice thing about this is it creates actionability too, right? If you clump everything together and you report that out it makes it a little bit more difficult.

And I can see in my own mind, and this is only a personal opinion, but cardiac rehab is so separate that the idea of lumping

it all and then not understanding which one is perhaps the gap that you're trying to deal with makes it more difficult.

The last thing I would just say is that the individual even though it totally mimics it, I think the fact that you're attributing at a different level is enormously different. It's important. And we as a group have decided that, again, it creates a different market which you can kind of have an actionable insight into quality improvement.

Most of that is editorial, obviously.

DR. KOTTKE: Any other comments? Yes, Tom.

DR. JAMES: My question has to do with attribution. And this is set up at the clinician level and it's coming from a specific data set right now.

Expansion of this type of thing though would have to recognize the matrix phenomena that goes on in hospitals. I was a

hospitalist in the past.

Who do you attribute this to when we're creating a measure?

DR. NALLAMOTHU: I mean, absolutely that's a very important point.

So, we've decided in the creation of this measure to focus on the interventional cardiologists who perform the PCI. If you had multiple PCIs it was the last person who performed the PCI during that hospital stay.

I think there's probably two reasons for it. The first is that we do feel that the interventional cardiologist after they perform the PCI in this patient population, they're very critical in terms of setting a lot of the mechanisms in place.

Even if it's not their responsibility, if it ends up being a cardiologist on the floor or some other care provider, I think the interventional cardiologist making that type of recommendation and pushing forward with it

plays a big role.

And the second is at a certain point to be practical about it. Exactly, exactly.

DR. KOTTKE: Liz.

MS. DELONG: I'm just a little confused about our role here. I mean, we endorse quality measures as a reflection of quality. Are we also responsible to assess the attribution? I mean, I would think that would be -- if it's being collected at one level it can certainly be decided by whoever's using it whether to break it into other levels.

DR. WINKLER: Typically I think that questions of attribution come up all the time in terms of specification. So, I think that question being addressed within the specification can be helpful. But you're right, the actual ultimate implementation program may make other determinations in terms of attribution.

Because I'm not aware that ACC is saying it's the PCI operator. I didn't see that ACC baked into the specification specifically. So, whether it's truly baked into it or not, or is it the way it's currently being used by ACC in the registry.

DR. KOTTKE: Any other -- Henry?

DR. TING: I just want to comment that everything we're saying about attribution and whether this reflects excellent care after PCI. All these comments ACC are referable to the prior one which you approved 100 percent.

Can we attribute using these medications at the hospital level yet we didn't have this discussion? Can we say that PCI, everything was done perfectly after PCI and was at the hospital level? So all the comments we've made so far are referable to the prior measure which we approved 100 percent.

DR. KOTTKE: Further -- Liz, are you still up? Yes.

DR. HOLLANDER: I don't want to dwell on it too much, but the attribution to the individual physician, as I think through it. We just approved a measure that gives it to the hospital, right? Now we're attributing to a physician who may not actually be the last provider which in essence is the same as the hospital. It's the care pathway for that individual patient.

And I can't speak to this as an ER doc, but it makes me think what do we learn if we're not 100 percent sure the meds are written at discharge by the person we're attributing it to over and above the measure we discussed before lunch?

DR. KOTTKE: Further comments? How about if I call the question here? Anybody want to change their vote on the composite construct?

DR. NALLAMOTHU: Can I make just one point? That's a great point and I think it's one that this group needs to take into

consideration.

I will mention that at least in this sample that we saw, and again with all the limitations I'm sure we're going to discuss, about half the providers practiced at one hospital. Then about 30 percent or so practiced at more than one hospital. And then about 20 percent practiced at more than two hospitals.

So there is in the modern practice of cardiology this idea that people do move around.

And that potentially has some implications for how you consider this attribution issue.

DR. KOTTKE: Henry.

DR. TING: Not to beat a dead horse, but the perfect attribution is actually at the patient level. But we don't have data to make a measure like that.

DR. KOTTKE: And it's usually N of one trials.

Okay, so unless somebody chooses

to change their vote from the prior vote we'll just use the prior vote. Thank you. Michael.

DR. CROUCH: Okay. As for reliability the only difference is the denominator specification. The exclusions listed differ in addition to patients that died, physicians who are discharged to hospice, or discharged to another acute care hospital, or who left AMA, against medical advice, are stipulated as inclusions in this one and not in the previous one.

And that's a harmonization issue that you may want to comment about. How do you want to plan to reconcile that.

DR. KOTTKE: Yes, Chuck.

DR. HOLLANDER: I was the secondary on this one. And I think that's an issue.

The other thing that wasn't addressed is the way they discussed in the last measure that if you had a contraindication to drug 3 you still counted

in the denominator for drugs 1 and 2. And so I think I'd like to see these two measures harmonized more precisely.

DR. NALLAMOTHU: You know, I'll give you the easy answer. This is one of the great things about the process of kind of vetting and going through this you guys are absolutely correct. The entire intent was to make them harmonized.

There were two areas and both of those we're working on right now. Part of the issue is that this measure has, you know, the ACC was responsible for the last one. PCPI was responsible for this one. And that caused a little bit of the issue.

But absolutely. I mean the last thing we want to do is create confusion around this. In fact, we might even think about harmonizing more the titles as well which would be a big issue. So we absolutely agree.

DR. WINKLER: Well, the title I think is a perfectly good one.

I think the other -- the exclusion exception issue I think was the other one the group brought up as being areas maybe in need of true harmonization as opposed to just writing the same words even though the intent was the same. So, to the degree we can clean up the things you truly are already the same.

The question is going forward what are the real differences between these measures and are they important differences that should continue. I mean, in all honesty true harmonization of these measures would make one measure go away. And it would be just multiple levels of analysis for a single measure.

So, the question is what are really the differences between the two measures and how does ACC see potentially going forward with true harmonization of these measures.

DR. KOTTKE: I think one of the issues is that the cardiologists really aren't

asking within hospitals. It's not like Mayo, you know, where you work at one site and if the hospital does great then every cardiologist in the practice does great. But there's the rovers and folks at multiple hospitals. And so they're not quite -- they don't overlap. They're not nested.

Are we ready to consider reliability? Yes, Sana.

DR. AL-KHATIB: I'm ACC totally in favor of what you just said because they are exactly the same. I mean, the only thing that was mentioned is in terms of how they worded the exclusion criteria. But beyond that the only difference is the level of attribution, the level of the measure.

And I would be totally in support of having, combining the two into one measure but having different levels of reporting.

DR. BURSTIN: Just one brief comment on the differences. And I assume this is correct but Jensen can correct me if I'm

wrong.

I assume part of the difference as well is because it is a PCPI-level measure. It has three fairly open-ended exceptions for medical reasons, social reasons and another reason I forgot.

But that does change the -- I mean that is enough potentially to make it -- the question is are those acceptable differences. Because again, remember under reliability, 2A1 here is precise specifications. So that should be a consideration for you.

MR. CHIU: Actually, as you know, we looked at the application, to Dr. Nallamothu's point. It is true. So PCPI did lead the effort for this measure. But the exclusions as specified actually should be identical to the NCDR ones.

But having said that though, for the other measures that isn't always the case. But for this we do have like, for example, the contraindicated. And how we calculate it in

our measure, how we noted it in the form, actually I think there are some discrepancies. So the calculation is exactly the same as the NCDR, number 0964 if I remember correctly, that measure is the same.

But I do think as we get to the other sections there might be some differences with usability and things that we'll talk about in a second.

But just to tackle the other question I think in terms of harmonization with cardiac rehab and things, I'm wondering -- Reva, I'll leave it to you. But if that would make more sense after we look at all the other individual measures before we do that. Because the cardiac rehab measure I know is not today but it's tomorrow.

Those two measures, that 0642 and 0643, actually there is PCI in there and that's harmonized across all the registries and everything. So there's heart failure, AMI and all those others. So I wonder if that

discussion might be helpful when we're looking at that measure specifically.

DR. WINKLER: Well, I just want to caution everybody that our role here isn't to make new measures. You can suggest things that might be measured and you would like to see measured instead but really we want to evaluate what's on the table in front of us.

So at this particular point I think the question to you is this is a new measure. What's its added value to the portfolio. And I think that there is the consideration of whether another new measure is necessary, or whether it can be incorporated into the existing measure or not, if they are truly identical.

DR. VIDOVICH: I just have one question. I think you answered it partially. We are not creating a new measure or harmonizing the measures. Is that correct?

DR. WINKLER: Ultimately the developers make any changes to the measure.

All we can do is recommend to them and base our recommendations for endorsement on our evaluation.

DR. VIDOVICH: But if this is the case I might just want to get the opinion of the group. I feel that the description "optimal medical therapy" might be a little bit too broad.

I think the aspirin, P2Y12 and statin is way more -- because optimal medical therapy is a large term. Brahm, as you mentioned, you can throw in beta blockers, ACE and ARB. So perhaps maybe limiting the scope of this measure. If you're harmonizing.

DR. KOTTKE: Any other discussion on reliability? Anybody choose to change their vote? Sana?

MS. LUONG: For the purpose of the people on the phone I'm going to say all the options. For reliability you can vote 1 for high, 2 for moderate, 3 for low and 4 for insufficient. And we can start the timer now.

For this criteria 3 voted for high, 13 voted for moderate, 3 voted for low and 2 voted for insufficient.

DR. KOTTKE: Validity testing. Michael?

DR. CROUCH: The validity issues are the same as for the hospital-level analysis are the same issue. I don't see any differences or significant issues there that are different from the others.

DR. KOTTKE: Any concerns? Anybody choose to change their vote? So, let's use the prior vote.

Feasibility?

DR. AL-KHATIB: So one point I brought up is the MPI issue. Because when they did the testing on validity they had a large degree of missingness in terms of the MPI. That's how you're going to attribute it to the physician. And when we brought this up during the call my understanding is that that was something that the developer was going to

look into to potentially ways by which you can minimize this large degree of missingness.

DR. NALLAMOTHU: We're going to let Lara do that.

MS. SLATTERY: Hi, Lara Slattery from ACC. So, as you see it takes a team to get a measure through your NQF endorsement process.

I should clarify that within the CathPCI registry for actually numerous versions we've had the ability to capture the MPI at the individual clinician level.

We have only recently begun using that data. And so what we know is that we did not spend a lot of time in earlier versions, or even earlier data reporting periods validating MPI that was inputted.

We recently, and I mean very recently have undertaken some mitigation steps. That started with outreach to the hospitals asking them to verify that they are entering in accurate MPIs for valid clinicians

that are performing the procedures.

We then have externally validated the MPIs that we've received from the hospitals up against the data that's available from -- that you can download from the government. And now have actually built the pathways that allow the physicians to access that data. That's a relatively recent activity. And we will continue to monitor that to see what additional mitigation we need to put into play.

For instance, if you know anything about the NCDR's registries and the data submission, data actually goes through some validation of completeness as well as validity of ranges in some instances. We have not taken steps to up that threshold or put in valid ranges for that but we may choose to do that moving forward.

So, it is relatively newer for us to be paying as close attention to the MPI. It is designed to support clinicians being

able to get access to that data.

And a lot of energy had to then be expended from a resourcing perspective on mapping it so the individual clinicians can now look in and view that data as well. So, it is an area that we are working on. It's a relatively recent effort.

MS. BRIGGS: I personally don't see that. It's a tough fix if you decide to fix it.

DR. NALLAMOTHU: I was just going to add like, you know, one of the funny anecdotes is 007 was apparently one of the most common MPI numbers. But in the last couple of years that's gone away.

DR. KOTTKE: So, validity. Anybody choose to change their vote? Nobody chooses to change the vote. We'll use the same count.

Again, 2d. Composite. Anybody discuss? Anybody change their vote? Seeing nobody changing their vote we'll take the

prior vote.

We're to feasibility, I believe.

DR. CROUCH: I don't think there are any different issues with this.

DR. KOTTKE: Anybody need to discuss?

DR. HOLLANDER: Yes, so you know, I'm now thinking about it. There's 10 to 20 percent of hospitals that aren't in the registry. And what if I'm a physician who participates at hospital A which is in the registry but hospital B doesn't. Is that going to give an accurate portrayal of my care pathways?

And so I don't know that feasibility is the right place for it but it is feasibility in measuring that individual physician. And I just thought about that. And I think that makes this a little different than the last measure.

DR. NALLAMOTHU: I mean, again, that's a great point. It does get to the

complexity of how physicians aren't necessarily nested within hospitals.

I think the only response I could really come up with, and again, understanding it's a great point, is that at least the care in those hospitals where that physician does participate and that are visible within the registry will be apparent.

You know, regardless of that care it's going to be the same issue as before that care, at least at this point in time. I mean, there are just a handful of hospitals that are out there but those hospitals are essentially invisible to these measures.

DR. KOTTKE: Any other concerns or comments. Anybody wish to change their vote on feasibility? Seeing no one. Okay, should we vote? You want to vote? Okay, let's vote.

MS. LUONG: For feasibility 1 is for high, 2 is for moderate, 3 is for low and 4 is for insufficient. And the timer will start now. Four voted high, fourteen voted

moderate, three voted low and one voted insufficient.

DR. KOTTKE: Usability and use. Anything new?

DR. CROUCH: I don't believe there are any significant differences between this and the hospital level.

DR. KOTTKE: Anybody care to comment on usability and use? Seeing no comments -- oh.

MS. STEARNS: Just quickly. From the perspective of consumers I think that it is not uncommon for report cards to reflect both hospital and physician information. So, consumers do often look at that information.

So, if in the end the data that is collected is identical that will be informative. But I think it's worth pursuing. Because you find out if there will be meaningful differences between whether the hospital-level data and the physician-level

data is the same. Because if there are meaningful differences among different physicians consumers if you're having elective PCI would want to know that.

DR. KOTTKE: That's true. Other comments? Anybody feel the need to change their vote on usability and use? Seeing no indication we'll use the prior vote.

We are to committee voting on whether to recommend measure for endorsement. Any discussion?

DR. AL-KHATIB: A quick question. If we end up endorsing this what will happen? I mean, you'll have these two measures, very, very similar. Not identical, I agree, but very similar. Do we really need to have these two measures in place?

DR. KOTTKE: Reva says that's the key question. I agree that your vote here, I mean if you vote yes to endorse this measure you're saying there's need for two measures. I believe there's a need for two measures.

Judd.

DR. HOLLANDER: So, I'm just a little confused. Because I know we're not supposed to reinvent measures. But we've sort of given advice and insights which the measure developers think are good ideas. Reva said something about oh, they could change this, they could change that. Is there like a "yes, but" vote? You know? So if I vote yes now does it mean the measure as is goes to the next step and it's never modified again. So do I need to vote no to get the modification so I could vote yes next time? And that sounds funny but it's a serious question.

DR. WINKLER: The question would be what's your modification. Let's talk about what it is you're actually talking about. I mean, are we talking about harmonization? Or are we talking about something else?

DR. HOLLANDER: So I'm talking about harmonization and other things raised here. But what if they go back and they

looked at some of the sort of low-hanging fruit that they said we can easily look at that and we don't think that's going to be a problem.

But it turns out they can't get MPI numbers on people. And does the measure then go away? So you know, they have a lot of good plans but they haven't proven they can do the things that we've asked to have fixed yet.

And we just had a sidewalk conversation about, well, what if physician A and hospital A is 98 percent but at hospital B they're 82 percent. You know, then it's really a hospital difference and not a physician difference. And they are looking at that but we don't know the results of that.

And so I think maybe I need to know the results of these things, maybe I don't. But if they find that they're exactly the same across all hospitals, well then I think the measure is really valid. If they find it's a crapshoot over all these different

hospitals for the same physician the measure is not valid. And we don't know that yet. So those are the kinds of things I'm talking about.

DR. WINKLER: I think at this point just because we'll use the same approach to all measures is you're voting on what's been submitted to you now, not the potentials for going forward.

Once we have the on this measure then we have the conversation about related/competing. If there are recommendations you want to make about further harmonization for the developer to take under advisement and hopefully maybe react to then that can be part of that secondary vote. But right now you're going to vote on what's submitted.

DR. KOTTKE: So, if they harmonize then it comes back for another vote here?

DR. WINKLER: You would see it back once the harmonization has occurred.

Because sometimes that's not something that happens within a matter of days or weeks.

And remember, you're a standing committee. That's what's going to facilitate them bringing things back. So, that's why you vote today on what's in front of you.

DR. KOTTKE: Yes, Sana and then Henry.

DR. AL-KHATIB: So let's assume the best case scenario, that they're able to convince us that the MPI data can be achieved and they're accurate, that they can harmonize it exactly with the other measure.

I guess my question that I still would struggle with is what is the added value from having this measure to the other one. If we have the ability to collect the information on the other one and report it based on different levels. I'm not sure I can see the added value from having this in our portfolio.

DR. CHO: I agree. And I think one of the things is that once these things

are endorsed then it's difficult to change them. And I think that right now there are so many moving parts in this current measure, the missing MPI numbers, the doctors going to two different hospitals, you eliminating physicians who do less than 50 PCIs a year. There's so many missing and moving targets that I just don't think that currently as this measure stands this is ready for prime-time.

DR. TING: So, for discussion purposes I would argue that this measure at the clinician level is useful. If you think about patient satisfaction you can think about it at the hospital level. But thinking about the individual clinician level as Christine says does give you additional information. Because it gives you a little more granularity about the individual clinicians.

And if you are hospital leadership or executive one of the best ways to engage your staff to do quality improvement is actually to report individual clinician-level

data as opposed to just hospital-level data.

Having said all that I also understand the comments that are being made which is if the prior measure could just be stratified at hospital, clinician and other levels then we wouldn't need this extra measure. But that's a strategic issue that's not what's in front of us and I'm not exactly sure I know how to deal with that.

DR. VIDOVICH: Just a quick comment. Physicians don't practice in a vacuum, right. You know, hospitals have systems of care. They have ACS order sets, PCI order sets and I feel it's tough to separate one from another. That's just my view from the two measures. So they probably would be better off to be harmonized and merged into one.

DR. KOTTKE: On the other hand, it's the physicians who do drive the order sets. I mean, I agree that context makes a huge impact. All of us that have practiced at

several different locations, we're different doctors in every place we practice. But it's we who drive the quality in those hospitals as acceptable. We accept it or we don't accept it.

Are we ready to vote? Yes, Liz.

MS. DELONG: I'm still confused. If this becomes harmonized it is one measure. It is one measure with two names. I'm not sure that makes sense.

DR. KOTTKE: I don't think it's one measure with two names. Because doctors aren't nested within hospitals.

DR. NALLAMOTHU: Can I make a comment? So, this is obviously a very interesting discussion. And I do hear a lot of the concerns. And I think it's very interesting to kind of hear this.

I would make a couple of points. I think the last point made by Christine here about usability, people do use these measures different at different levels.

The second is the one that I've continued to struggle with which is what Tom has mentioned multiple times is if you do just create this at a different level of attribution is it the hospital that's just going to aggregate within their own group what their operators are doing and each of the different hospitals is responsible for that. And you never get a cross-institutional view.

And then the third thing is, you know, maybe we've been thinking about it naively, but like Judd has mentioned which is this question, and we did have this sidebar conversation.

But you know, we see it as important regardless. So if there's consistency across hospitals that tells us something about the operators being involved.

But if there is inconsistency across hospitals while it does get at the hospital being responsible more so there is no more important lever for like actual clinical

action than to have an interventional cardiologist not do well at a visible way.

And so we think it's important but we're not sure if it really matters for this measure in general. And that's kind of how we thought about it. So.

DR. KOTTKE: So I think it's time to call the question. So, if you vote yes on this you are -- the measure would be as stands. You could vote no meaning that they should harmonize, change the title, et cetera, and come back and -- or you could be voting no because you think you don't need another measure.

DR. BURSTIN: One clarification. So harmonize means like measures are actually harmonized. They have the same characteristics that fits here.

What you're talking about going beyond that is saying it's one measure with different levels of attribution. I think that's what people are struggling with.

So I think I heard Jensen say that any of the discrepancies are unintentional and they will in fact be fully harmonized. Is that correct, Jensen? Across the two measures.

MR. CHIU: For this measure that is correct. I know another one coming up is a separate issue. But for this one, the exclusions, I know there are some issues in the application. Those exclusions and exceptions are harmonized.

DR. BURSTIN: So these measures are actually fully harmonized or will be fully harmonized by the time they come back to you.

And so the real question is is there a reason to have two measures or one. And I think you just heard the discussion of how you get a different population when you look at this versus hospital because you may just get physician cluster within the hospital. I just want to be careful with that language. Because in fact they're telling us

they will be fully harmonized. They just may be two instead of one to capture both levels of analysis.

DR. KOTTKE: So who decides that they're harmonized. Is that you, Reva? Is that NQF? I mean does NQF say --

DR. WINKLER: I think we're basically listening to what ACC is telling us about the measures just as all the information about the measures comes from them.

So indeed, what I heard is the fact that even though there may seem to be differences in the written materials in fact that was not meant to be and that they should be essentially identical.

DR. KOTTKE: So, do we as a committee look at it again and give it final approval? I mean, is this a "yes, but" vote?

DR. BURSTIN: It could be if that is something we need to do. We can take a look at Jensen sends us back. If it's literally identical with the exception of

where it says level of analysis then we can probably just share that with you in an email. But we can clear that up post hoc. Right, Reva?

DR. KOTTKE: Does everybody understand what they're voting on?

DR. AL-KHATIB: No, I'm not sure that I do. So does this mean that it will be one measure but you have different levels of reporting? Or it will be two different measures? With the only difference being the level of reporting.

DR. BURSTIN: The latter. Because I think what they're telling you is that if it was a hospital -- if they just added a level of analysis it would be nested within the hospital is I think what I was getting from you. As opposed to the fact that physicians can be across multiple hospitals.

DR. AL-KHATIB: But if the analysis is done when using the MPI how does that not capture the procedures that you do at

different hospitals?

DR. KOTTKE: It does.

DR. AL-KHATIB: Right.

DR. KOTTKE: But if it's at the hospital level you only capture a portion of the --

DR. AL-KHATIB: So I guess what I'm not clear on is what is the added value of having the two measures if we just go with the initial measure that we all endorsed and say let's report it at different levels. Report it at the level of the hospital. Give the option of people to report it at the level of the healthcare provider. And they would use the MPI and that would capture all the procedures that that provider does regardless of whether they're doing them.

MS. SLATTERY: Lara Slattery again. I just want to clarify that while a lot of the responses may appear to be ACC only responding in fact this is a different group putting forward this measure for stewardship.

So, in the previous measure it was only ACC that is being put forth as the steward of that measure for implementation which includes a lot of decisions around usability for that measure.

In this instance this was developed as a PCP/ACC/AHA measure. ACC/AHA will take over stewardship of it. And so that does change -- that's the only mechanism by which we can find to submit the measure. So they are in fact two separate measures in part because stewardship of those measures is governed differently.

DR. KOTTKE: Yes, Liz.

MS. DELONG: We now have two measures that are presumably harmonized but overseen by different groups. But it's the same measure nonetheless. It is described exactly the same way. And are we at risk of expanding this portfolio to be uncomprehensible?

DR. KOTTKE: I think people have

to decide whether it is the same measure for themselves. Judd?

DR. HOLLANDER: So I think we're measuring the same thing but we're reporting different things. And I kind of think it's the lumper and splitter argument, whether you call it one measure.

If there's going to be two voluntary reporting websites, one by the physician and one by the hospital, then I'm fine either way, whether it's one measure or two measures because you're filling out the same data set in the same registry going to the same place.

And so, I don't know, it doesn't matter to me if it's a different title on a different website, or it's a subcategory of the first website. So I'm okay with it as a second measure because I think it's really the same thing.

The amount of work on the hospital end is going to be the same as one measure

rather than two. My biggest concern is that I want to make sure they get harmonized and I don't know if I give the "yes, but" number 3 in order to do that following the rules of NQF.

DR. KOTTKE: My understanding is it would come back for a final vote to us, maybe an email vote to prove the harmonization. I think it's time to vote on this very straightforward issue here.

DR. CROUCH: Can I just make one last comment? As a family physician who sends patients to cardiologists all the time I'd like to see the cardiologist data reported by individuals rather than hospital. And I'd like to have that data be available sooner rather than down the line.

DR. KOTTKE: I think Christine's comment that patients would like that too.

Okay, it's time to vote. 1 is yes, 2 is no. Vote your conscious.

MS. LUONG: The timer starts. So,

11 voted yes and 11 voted no.

DR. WINKLER: I think this is a perfect example of consensus not reached. It is.

I think that perhaps given the conversation we've had this will be an opportunity to allow ACC to verify the harmonization.

Also, we can put it out for comment with consensus not reached and see what the world out there wants to tell you and bring it back for another review for you all. Does that seem like a plan?

DR. KOTTKE: Yes, there's clearly considerable interest in this and it's around the measure. Encourage ACC to clean it up, bring it back. Henry?

DR. TING: Can I just make one comment about process? Because if this measure had been reviewed first instead of the other one it could have been very different. And I'm not sure this process is equitable to

this measure compared to the other one we just reviewed and approved 100 percent to zero.

DR. NALLAMOTHU: And I have to say one other thing too just to build on that is that, you know, I found it interesting to go through the entire process. And then, I didn't know at the end whether you were going to accumulate what you had done.

But this reminds me a lot of study section, right? Everybody breaks down different things and then you're like all right, well, where did you get the impact score at the end of the day.

So, just -- and I'm only mentioning that because from the measure development side, I mean we would want guidance as to where we fell short in this particular regard. And so I think that would be an important charge for you guys.

DR. BURSTIN: And I would suggest that before we put this out for comment we allow ACC to go back with PCPI and kind of

work this through. I think they just need to kind of work it out amongst themselves.

You're absolutely right. Henry, there's absolutely nothing about this measure versus that measure. It's just that clearly half of you don't want two of them.

So, please go back and we'll figure it out to follow. We can do it in email.

DR. KOTTKE: Would it be appropriate to get sort of a hand vote on how many people think there ought to be just one measure?

DR. BURSTIN: Is that what that was?

DR. KOTTKE: No, I don't think so. I mean, there's a whole bunch of questions in there about harmonization and title. How many people think that this should be rolled --

MS. STEARNS: Is that possible? Do we have measures where we measure both hospital-level data and physician-level data?

So that happened. Okay.

DR. HOLLANDER: And you combine them across hospitals. Like, the advantage of this measure is -- you can do that. Okay.

DR. KOTTKE: But if only hospitals are reporting then you don't have -- you don't really know how the cardiologists are doing.

DR. SPANGLER: I have a question for Reva and Helen. I mean, this is a process question. Because if you look at the voting up to this point it met all the criteria to be endorsed. But despite that many people voted no even though they voted that it met the criteria.

So, does that mean -- I know that's happened before, but the question is are we missing something then in the criteria?

DR. KOTTKE: It has to do with composites.

DR. WINKLER: No, I think that you combined really two votes. One was suitability for endorsement as well as what we

would have -- you would go into the next question which is the related and competing issue. Because your vote on suitability for endorsement wasn't final pending the discussion of related and competing measures which you kind of pushed together.

DR. AL-KHATIB: So what I wanted to add is exactly that. I mean, all of us actually like this measure but we still don't see the added value from having it as a separate measure, knowing that the first measure can actually be reported at different levels. That's the missing point for me anyway.

MS. MITCHELL: I think the issue really comes down to we were asked to vote on what is on this piece of paper right now, period.

And I think as a part of the process we discussed what it could look like. And I think there was opportunity to conflate could with should and is.

And so I think going forward just keeping in mind that we're supposed to be talking about what has been submitted for review for endorsement today. If that's incorrect please let me know but that's how I'm operating.

DR. KOTTKE: I hate to have ACC work on this whole thing and have it rejected again. How many people would like to see this come back cleaned up? Just a show of hands.

A separate measure that they feel that ACC's time is well spent to harmonize it. It comes back as a second measure. Maybe the title is changed so it's not quite as broad, that was brought up. Combining it to one measure with the other measure. So they work on it, come back. So there's two measures, there's a hospital-level measure, there's a clinician-level measure, they're harmonized at all aspects except one is hospital, one is physician. I'm the only one?

MS. TIGHE: And I do think we need

to clarify. It's not necessarily that these are ideas in opposition to each other. We don't know that ACC can expand the level of analysis for the first measure. So it may be that we have two measures that measure the same thing at different levels of analysis because they have some stewardship issues. So I don't know if you guys want to speak to that.

MS. SLATTERY: Yes, I mean -- again, Lara Slattery. I do want to emphasize that this is -- this measure being put forth is a collaborative measure that is jointly developed with our partners the American Heart Association.

So, you know, I appreciate and our desire is to have a harmonized measure that is efficiently leveraging the same data source that is accurately reflecting to the best degree that we can the performance of the clinicians, understanding that they may not have control over the data being submitted

because they don't directly make the decision of whether to participate in the registry or not.

However, if the recommendation is to create one measure it is from our perspective somewhat disingenuous to the contributions that our partner societies, in this instance PCPI and American Heart Association have made in developing this measure.

So I just don't know how within NQF's structure we can reflect those stewards the way they would like to be acknowledged in contributing to this measure which is why you have two measures that have been put forward.

DR. BURSTIN: And we can certainly work with you on that. I mean, actually, Mary probably knows this best from the stroke world how many co-stewards there are, for example, on the stroke measures. That's not a problem. There's a way to in fact make it ACC/AHA/PCPI for the combined measure. We can work with

you on that.

MS. SLATTERY: But the reverse may not be the case where they want to accept stewardship at the hospital level.

DR. BURSTIN: Well, they can be a co-developer but not the steward. There's plenty of -- I mean, don't let those technical legal issues affect what you think is the best way to get the measure information from docs, hospitals and get the best information out there.

MS. SLATTERY: So then in essence these are the same measure, it's just --

DR. BURSTIN: Yes.

DR. KOTTKE: Tom?

DR. JAMES: I don't know whether you want to invite more comments and I can shut up if that's the case. But it seems that I've grabbed the floor.

Rob Huckman at Harvard has made the point that if there's not a significant variation among physicians in an area, that

perhaps that's not a good measure to look at. It's better to look at whether it's the variation in the therapies offered.

In this case when I look at this data, the difference between the 75th percentile and the 25th percentile is not that great. So to me I think this is a better hospital measure than a physician differentiator.

MS. TIGHE: On that just to clarify process. So, when we draft the report we'll post it for NQF member and public comment. And that will give ACC some time to consider these issues that you've raised and potential responses to them.

We have a call after the comment period where you'll consider all of the comments, any additional information from ACC, and you'll have the opportunity to re-vote on the measure at that point in time. So, this is a first vote but not necessarily a final vote.

DR. KOTTKE: I think it's time to move on. Thank you.

DR. GEORGE: So, just to let you all know we're a little bit behind schedule. We'll be going to the next measure, adherence to antiplatelet therapy. Are the developers available?

MR. CAMPBELL: Hey, Reva. This is Kyle Campbell at FMQAI. Can you hear me okay?

DR. GEORGE: Yes.

MR. CAMPBELL: Okay. Did you want me to kick off the measure?

DR. WINKLER: Yes, Kyle. Go ahead.

MR. CAMPBELL: Thank you. All right, well, good afternoon. My name is Kyle Campbell and I'm the pharmacist and executive director at FMQAI for the CMS Medication Measures Special Innovation Project. Our project is tasked with both maintaining and developing new medication-related measures for CMS.

The measure submitted for your consideration today really picks up from the prior measures and focuses on adherence to antiplatelet or P2Y12 inhibitor therapy for patients in the 12-month period following stent placement.

As directed by NQF we worked closely with the Pharmacy Quality Alliance to establish a standard methodology for NQF adherence measures. And the PDC methodology or proportion of days covered methodology selected was based on extensive testing to establish its validity.

The measure was developed under the guidance of a multidisciplinary technical expert panel and has undergone rigorous development and testing processes as specified by the CMS measure management system blueprint.

The measure is based on administrative claims data and has been tested with 100 percent 10-state sample and also a

convenient sample of 31 accountable care organizations.

From an importance perspective this measure addresses two of the National Quality Strategy goals, namely promoting effective treatment practices for the leading causes of mortality and also engaging patients in their care.

Stent placement procedures are frequently performed. They account for high resource use and lack of antiplatelet adherence is associated with severe patient and societal consequences.

As this is a shared accountability measure we are proposing the measure for multiple levels starting with the physician group, moving up to health plan and accountable care organization as well as the state level.

Finally, we did receive questions in our workgroup review of the measure from the steering committee. And we have submitted

a memo under separate cover answering those questions as requested.

We appreciate your consideration of this measure today and look forward to answering any questions you may have. Thanks.

DR. GEORGE: Thank you. And we'll move onto the primary discussant.

MR. BURTON: Yes, hi, this is Jeff Burton. Can you guys hear me okay?

DR. GEORGE: Yes.

MR. BURTON: So, since Kyle gave that very detailed introduction I'll hop right into the evidence.

Obviously this is a process measure that demonstrates medication adherence and how it potentially leads to decreased adverse cardiac events and lower mortality rates.

The overall body of evidence is good when it comes to supporting the use of antiplatelet medication following a PCI. I don't think many would argue that.

Some of the intricacies I think of how we actually measure adherence to a medication is where we may run into a couple of challenges that were noted during our workgroup call and that Kyle provided some clarification or some answers to.

So, to give a brief overview there were three practice guidelines that were presented. They did not have QQC ratings but they were important to establish the guidelines for the use of antiplatelet therapy following a bare-metal stent or drug-eluting stent, all of which were class 1 level A or B recommendations.

The one thing here to note though is that the guidelines, one of the guidelines for bare-metal stents in non-acute coronary syndrome did indicate that clopidogrel be given for a minimum of one month and ideally up to 12 months.

I think that in the response to this that it only represents about 67 percent

of the members in the denominator. And that the technical expert panel made a recommendation to include these patients in the denominator even though the evidence wasn't definitive on a time period and stated that it was superior to have the therapy for 12 months as indicated in the measure even though that the body of evidence said that 1 month as a minimum would be sufficient.

There was a systematic review providing evidence that related directly to actually adherence of medication by a discontinuation of clopidogrel at different points following the stent.

The QQC for this was high in quantity, moderate in quality and one could argue low to moderate in consistency as some of the studies did have different directions that supported the data.

So, two additional studies were conducted where a critical threshold of 80 percent medication adherence was established

given the difference in mortality rates for cohorts that had below or above 80 percent. So taking all that into account and using the NQF algorithm to rate the body of evidence I believe it could fall into a moderate category. And I'll leave it up for the committee for discussion.

DR. GEORGE: Do we have discussion on the evidence for this?

DR. HOLLANDER: I sort of have a problem with this one because they're using the term "adherence" and none of this is about adherence. It's about did the medication get filled.

And so if you're in a prescription plan where every month or three months they send you a 90-day supply and you never take the medication it appears to be adherence. And so I think it's sort of a fallacy here that it just depends on your prescription plan as to whether or not you're going to appear to be adherent. So I don't think they're

actually measuring what they claim to be measuring, at least the way I read it.

MR. BURTON: That's something that I was going to bring up in the usability of this. I know that medication adherence is very hard to measure because with the administrative claims data you're measuring prescriptions that were actually filled. And not so much the actual adherence of a patient taking those medications which can apply to any medication adherence measure.

I do know that the NQF has endorsed other measures relating to medication adherence based on administrative claims. Is that correct?

DR. WINKLER: Yes, it is. In fact, a couple of years ago we did have a project around medication and this was a huge issue, it was measuring adherence.

I would just ask the question, say the measure we just looked at where it was was it prescribed on discharge. Do we know the

patients ever took them there either.

I think it's probably the question that comes up most commonly with any measure around medication is it's a little hard to measure whether they put it in their mouth or not.

DR. HOLLANDER: All right, so I could see doing it at the ACO level or at the payer level. Because if we're encouraging payers to find ways to get medications into patients' hands it makes sense.

But it's hard for me to envision doing this at the clinician or institution level since they don't necessarily control all the difficult prescription plans the patients are on. And I think a lot of it will be driven by that.

MR. CAMPBELL: So, this is Kyle Campbell for the measure developer. Just a couple of points.

I think we aren't recommending this measure for the individual clinician

level. We are recommending it, however, at the physician group level. So if there's a group practice they can by using the data available from the measure be able to determine what the overall adherence pattern looks like in terms of fills for their patient.

MR. BURTON: Kyle, this is Jeff, primary discussant.

I know a couple of other committee members had some questions as to the amount of physician groups that were actually included due to the fact that there wasn't enough data. It wasn't reliable enough and there was only 13 percent of those physician groups.

So, we're jumping ahead here I know a little bit to the I believe feasibility. If you're going to be measuring at a physician group and you're only looking at about 13 percent of all physician groups that are able to have enough data to do the measure. And that's I think a little bit of

a concern.

DR. WINKLER: Guys, it would be helpful if we could right now just focus on evidence. It would kind of keep the conversation a little bit crisper for everybody.

MS. BRIGGS: So, we did talk a little bit about the fact that there is not sufficient data for bare-metal stent use of the P2Y12 inhibitor for 12 months. The recommendation within the guideline is 1 month to 12 months. And there's only evidence for that level of recommendation within the guideline.

So, the evidence really doesn't follow basically what's being asked for by this measure. The measure is basically blanketly saying anybody that got a stent should have 12 months of P2Y12 therapy. While that might be optimal that's not what the guideline says. And we were using the guideline as our evidence, then we're really

not following that evidence.

DR. GEORGE: Any other comments? Yes?

DR. VIDOVICH: My feeling is the measure may not completely accurately discriminate the acute coronary syndrome from elective PCI. Because then the guidelines change for 1 month to 12 months. As written it might cause some confusion because of this similar topic that you mentioned about the 12 months.

MS. BRIGGS: This guideline is only for electives.

DR. VIDOVICH: It's elective only?

MS. BRIGGS: Yes. This measure is only for electives.

DR. VIDOVICH: Okay.

DR. GEORGE: This is a really important point to consider when we look at these things right off the top. Any other comments?

MR. BURTON: I did have another

comment that I briefly mentioned. There were a few studies in the systematic review that they didn't show the same effect of clopidogrel cessation on stent thrombosis as they saw in other studies. So, the lack of consistency of those studies was a concern to me.

MR. CAMPBELL: This is Kyle Campbell again for the measure developer. I would just suggest that the additional studies did show consistency.

We do recognize that the -- for the recommendation related to the bare-metal stent for those non-acute coronary syndrome indication as has been discussed it was suggested that it would be optimal for 12 months of therapy. And when the measure was specified that was felt to be the way to go in terms of aligning everything with the ACC guidelines.

That said, since that time and since the workgroup we have looked at the

feasibility of stratifying by ACS and non-ACS. And we are able to do that.

And with the ACO sample there's about 2,000 patients overall in that denominator. And if you exclude patients with bare-metal stents for non-ACS indications that's about 10 percent.

The reliability of the measure does not change. The rate of the measure increase slightly from an overall mean of 0.78 to 0.80. And the range of the measure -- it still has a wide array of variation with a min of 0.69 to a high of 0.85.

DR. GEORGE: Do we feel we're ready to vote on this in terms of the evidence? Okay, we'll go ahead and vote.

MS. LUONG: So, for those on the phone 1 is high, 2 is moderate, 3 is low, 4 is insufficient evidence with exception and 5 is insufficient evidence. The timer starts now.

For evidence 2 voted high, 11 voted moderate, 5 voted low and 4 voted

insufficient evidence.

MS. TIGHE: So this just falls within our consensus not reached criteria. So we'll move forward with discussion of the measure.

MR. BURTON: So the gap in care, the opportunity I stated before, the critical value of performance was 80 percent for medication adherence. The developer evaluated performance based on the Medicare claims for eight states over a two-year period looking at the prescription drug plan level, looking at the state level, the physician group level and the ACO level.

The states, the plans and the physician groups all had -- or each had an average performance level of 75 percent, but the ACOs had a 78 percent. So there's a small gap from the 80 percent critical value of performance.

However, as a process measure ideally you get to 100 percent performance.

So I do think that there is a gap in care here and an opportunity for improvement.

DR. GEORGE: Discussion on opportunities for improvement? Yes?

MS. DELONG: I didn't follow where they got the data. If they can't measure adherence in a lot of situations where did these data come from?

MR. CAMPBELL: This is Kyle Campbell again for the measure developers. So, these data are derived from Medicare administrative claims data that include Part A which is generally the hospital, Part B which is the outpatient benefit and Part D which is the prescription drug benefit.

The numerator compliance is measured with the days covered from those prescription drug claims. So those data are readily available to calculate for the measure for this population.

MS. DELONG: So you can tell of the numbers prescribed which prescriptions

were filled and for how long?

MR. CAMPBELL: That's correct. We can tell which medication was filled and the days supply for that medication. And then that gets put into the measure algorithm to develop a days covered which would actually adjust slightly to the overlap of any fills in prescriptions.

DR. AL-KHATIB: Just a quick question. As we all know, a lot of the beneficiaries have other ways to get their medications other than CMS.

So, do you have a handle on what percentage of patients at least in your sample that you looked at had other ways, other coverage if you will for their medications?

MR. CAMPBELL: So, we did just briefly look at that with a sensitivity analysis where we looked at the potential frequency by imputing patients didn't have Part D-covered drugs, what would be the effect if we imputed 100 percent adherence rate on

those patients. And we didn't really find any effect.

And I will say that there's probably more concern -- even though this is also limited, there's more concern for drugs that would be on a generic formulary where patients would be likely to pay cash. In this case, you know, I don't think that that would be the case with any of the P2Y12.

So, it is conceivable that patients within our population could have a VA benefit let's say. But that would be true of all other NQF-endorsed adherence measures that are based on claims of which we're a steward of and any other organization is a steward of.

So, we haven't looked at it as a limitation particularly when there's a gap in care. And we know that as was said that these measure rates should be much closer to 100 percent. And we don't really think that that would have -- it would have a meaningful impact on the measure.

MR. MATTKE: And one more comment. Soeren Mattke for the developers.

Remember that in order to get identified for the measure we must see prescription fills under your Part D benefit. So it would only be of concern if people use sometimes Part D, sometimes other sources of coverage.

DR. GEORGE: Linda?

MS. BRIGGS: When we discussed this within the workgroup we did have questions to go back to the developer related to the fact that there is some gap in coverage in the Part D Medicare benefit. When patients get to a certain dollar amount they fall into the "doughnut hole."

Now, based on that information there could be potentially a gap which patients are supposed to submit the charges for those drugs so that they get credit for it so they get out of the doughnut hole.

However, depending on the

patient's other medications the timing of when that occurs is variable. So that if somebody was close to the end of the year, let's say November, and they just hit the doughnut hole, they may not be inclined to submit that data. So that the data set that they're working from is not perfect.

But just to point out that there are some reasons why patients might have adherence discrepancies that are not truly reflective of the patient taking or not taking the drug.

MR. CAMPBELL: So this is Kyle Campbell for the developer again. And we did submit a response to that question in a memo under separate cover on April 17.

Just a couple of points about that. CMS does require Part D plans to process claims and track the true out-of-pocket costs paid by the beneficiary in realtime.

Secondly, and I think maybe more

importantly is with the passage of the Affordable Care Act the Medicare drug coverage gap affectionately known as the doughnut hole will be phased out completely by 2020. And based on the current provisions within the act the amount beneficiaries pay for those out-of-pocket prescription drugs has already begun to decrease.

Originally it was 100 percent for both brand name and generic drugs in 2010. It's now for 2014 47.5 percent for brand name drugs and 72 percent for generic drugs. So there is an incentive for beneficiaries to have these claims under their plan.

And by 2020 the percentage will be 25 percent for all drugs which is essentially the same as the percentage paid by beneficiaries for up to the point of the coverage gap.

So therefore we anticipate minimal to no impact on the measure rates. This measure is new. It hasn't been proposed -- I

mean, implemented into a program so presumably it would be at least another year before it could be implemented in which case the Affordable Care Act would decrease even further the amount the beneficiaries pay in the coverage gap.

DR. JAMES: And it's just for those particular comments that have just been raised that I think this is a good measure for health plans and for large populations.

It becomes problematic at the smaller individual group level. But for a health plan it means I'm holding myself accountable. I think this is a fair measure.

DR. GEORGE: Are we ready to move to a vote?

MR. BURTON: I think so.

DR. GEORGE: On opportunity for improvement.

MS. LUONG: So, 1 is for high, 2 is for moderate, 3 is for low and 4 is insufficient. The timer starts now.

Eight voted for high, 12 voted for moderate, 1 voted for low and 1 voted for insufficient.

DR. GEORGE: Onto priority.

MR. BURTON: So the priority. Same thing as before when we were talking about the nature of the PCI and either medication following a PCI or in this case adherence to a high-priority given the sheer number of PCIs, given the cost per PCI.

But maybe even more importantly the importance of making sure that the medical community is focused on strong adherence in any way possible for their patients when things may be out of their hands just because something may be an imperfection in the measure and we should de-prioritize it as an important part of the software.

DR. GEORGE: Any comments on priority? Should we move to a vote on priority?

MS. LUONG: For priority 1 is for

high, 2 is for moderate, 3 is for low and 4 is for insufficient. The timer starts now.

If you could just keep pressing your vote here. Sorry. Eleven voted high, ten voted moderate, and one voted insufficient.

MR. BURTON: Maybe you could move onto scientific acceptability.

DR. GEORGE: Just one question. We're almost at 3 o'clock. Do you want us to start the discussion?

DR. WINKLER: Yes, let's go ahead and do that. But we will want to take a break shortly for public comment. Go ahead, Jeff.

MR. BURTON: Oh, okay. So, as far as the scientific acceptability again we're using administrative claims. The numerator is equal to the sum of the days covered by the days supply of all antiplatelet prescriptions during the days measured in the denominator.

The denominator is equal to the sum of the days measured for all individuals

who undergo coronary artery drug-eluting stent or bare-metal stent at any time during the first 12 months of the 24-month measurement period and have at least two prescriptions for antiplatelet therapy during the 12 months following the stent.

I think the key thing here is the two prescriptions at a minimum to capture those who may have intolerance or allergic reaction to medications which would throw them out of the denominator.

As far as any other coding issues the developer did submit a list of all the NDC codes as well as the contraindications which focus on intracranial hemorrhage, GI bleed and peptic ulcer disease.

DR. GEORGE: Any discussion? Ellen?

MS. HILLEGASS: I think I may not be able to find the information that was said to us before, but I was looking for an exclusion of acute MI. And I don't see it

anywhere in there.

From what I'm understanding the developer believes that this is for just new PCI, no AMI before. But I can't find this in the writing anywhere. Can anybody address? I have not been able to find it in exclusions. I haven't been able to find it in numerator or denominator.

MR. CAMPBELL: That's correct. We do not exclude those patients with the prior MI. And Soeren, I don't know, from RAND if you want to comment on that?

MR. MATTKE: Someone else might actually be a better person to comment on that. Can you clarify why we would exclude patients with prior MI?

DR. KOTTKE: Ellen was saying that she didn't find exclusion for patients with an acute MI, not prior MI.

MR. MATTKE: Oh. But we have patients with implantation for acute coronary syndromes which does include AMI and patients

with elective implantation.

DR. VIDOVICH: Did we mention that this measure excluded ACS? I was just told that a few minutes ago. Because I just searched like "eligible" through the document. I can't find that word anywhere in the document.

MR. CAMPBELL: The measure does not exclude those patients with ACS. It is inclusive of patients with ACS.

MR. MATTKE: Because the patients with acute coronary syndromes like unstable angina or acute infarction actually have a much higher risk for stent complications. So we definitely want to keep those.

DR. VIDOVICH: But the indication is for duration of dual antiplatelet therapy are different for elective PCI and ACS. Right? Hypothetically, pre-operative. I'd say pre-op BMS could get away with one month of dual antiplatelet therapy.

MR. MATTKE: No, I think the

recommendation is --

DR. VIDOVICH: ACS is 12 months regardless of the stent type. But non-ACS, they do differ.

DR. PHILIPPIDES: Right, but I don't think the inclusion of ACS would change that. You'd still have to give them dual antiplatelet therapy out for a year.

DR. VIDOVICH: But I believe that they should score non-ACS. Then the measure might incorrectly measure that they should have received 12 months whereas only 1 month might have been sufficient.

DR. PHILIPPIDES: Correct, but the problem here is not including the MI patients. The problem is requiring that BMS stable patients get 12 months. That's where their issue is.

DR. VIDOVICH: Correct. That's right.

DR. PHILIPPIDES: The MI is not the one that's --

DR. VIDOVICH: Yes, the MI is not a problem.

DR. PHILIPPIDES: It's the other guys.

DR. KOTTKE: But this would be a case where you could include both in a single measure because you have a different code. I assume that interventionalists code ACS differently than stable coronary. Yes. So here you could -- you put them both in the same measure.

MR. MATTKE: So, to go back. The measure does include both stable and acute coronary syndromes. The indicate, the recommendation is to treat all patients regardless of the indication and regardless of stent type for 12 months.

However, since the risk-benefit rate for stable patients on bare-metal stents is a little bit less favorable the guideline suggests that you could get away with at a minimum one month treatment.

To keep in mind, however, the way that you get into the denominator for the measure is that you have to have two fills which indicates to us that somebody is actually trying to treat the patient for longer than a month because the fill is actually 30 days.

So our assumption is once you get into the denominator it's the stated intent of the clinician to actually treat for a year because the risk-benefit rate has been determined to warrant ongoing treatment.

DR. TING: That's not completely accurate. Just to quote the guidelines it actually says two weeks for bare-metal stents in non-ACS patients. If there's a tradeoff for bleeding and risk of bleeding.

MR. MATTKE: Yes, but you can still see once you are in the denominator you must have been on 60 days of treatment already. So those were really -- it's unlikely that we are talking about patients at

that point in whom bleeding complications are a major concern because they would never be on 60 days to begin with.

DR. TING: There's probably a group of patients that you discontinue the DAPT because of upcoming cardiac surgery at two weeks.

DR. VIDOVICH: ACS in particular is an example right there where you have to discontinue because of delivery. Or upcoming surgery.

DR. HOLLANDER: I think the measure developer's point is that you have to get two prescription refills. So you wouldn't have gotten two refills if you're going to get stopped at two weeks, or if you're going to get CABG within the next month. So I'm still not sure I agree with that as the criteria but I think that's --

DR. GEORGE: Linda?

MS. BRIGGS: Again we come back to bare-metal stent recommendation, either two

weeks if there's a bleeding concern but one month is the recommendation, at least one month and up to a year.

A clinician might decide that the risks outweigh the benefits beyond a certain point in time for that particular patient and they would be totally justified according to the guidelines of stopping it even after two prescriptions. So it might be two months, it might be six months in that maybe it's an elderly person who has a fall and has some kind of complication related to that. There are a million reasons why a clinician might feel justified for that. And they would be well within the guideline parameters.

DR. GEORGE: So I'm hearing a lot of concern about the fact that both bare-metal and drug-eluting stents are included in this. Is that?

DR. PHILIPPIDES: The same recommendation for length of dual antiplatelet therapy. I think if they had tweaked it and

said for bare-metal stents we're going to really come out to about a month I think most of us would be okay with that.

But oftentimes there's nothing wrong with putting a patient on a bare metal stent and putting them on dual therapy for two or three months until they see you again.

Then you say you know, I've tweaked the drugs long enough. There was a reason I put a bare-metal stent in the first place. I was worried about bleeding. I'm going to stop it now. And that would be considered by the guidelines -- Henry, I think you'd agree -- perfectly adequate therapy.

The way this metric would have it was not adequate or wasn't as good as the other clinician. So I think that's what's giving us pause.

DR. AL-KHATIB: I completely agree with that comment.

The other question that I would raise is since we're using administrative

claims data I'm not aware of any way by which just based on the coding we can capture whether a bare-metal stent was used versus a drug-eluting stent. And without being able to make that distinction you either have to limit this to one type which you won't be able to capture. That raises certainly concerns about how we're going to be able to implement this measure.

MR. MATTKE: Soeren Mattke for the developers again. These are actually two different CPT codes. Partly because the drug-eluting stents are considerably more expensive. And so you can distinguish them in administrative data.

DR. VIDOVICH: I have a nomenclature semantic question. We are asking adherence. How do we know which duration was prescribed to the patient? Do we know that the patient should have received the month or 12 months?

If you're calling this adherence,

right? Because adherence would imply that we did know what the duration of therapy was prescribed. So how would you know this from this measure?

MR. CAMPBELL: This is Kyle Campbell for the developer. It's measured just the same way that all the other adherence measures are. We don't have specifically the ability to know the intent of the physician from the administrative data that they intended for 6 months or 12 months.

But we can see all the prescriptions filled and the days covered. And so those are basically added up to determine the proportion of days covered.

And in this case there's a fixed follow-up time such that it's one year post the stent placement after the successful fill of two prescriptions.

Just one more note. We have been able to operational because we did look at this after the workgroup concerns. We are

able to separate bare-metal stents and drug-eluting stents as well as determine from the claims data who has it for acute coronary syndrome and who has elective. So we can do that as well.

DR. AL-KHATIB: Actually, a question pertinent to this last comment. Have you done any studies to validate the accuracy of these codes in terms of using them for bare-metal stent versus drug-eluting stent or I think the other would be easier.

But especially in relation to this particular issue do we have any data that show that you have validated those codes and they're actually accurate?

MR. CAMPBELL: So, this is Kyle Campbell for the developer again. We have not done any sort of validation with the chart review to take a look at those codes. And I don't know, Soeren, if you have anything to add with regard to that.

MR. MATTKE: No, but it's unusual

to validate the coding accuracy because these are administrative data that get routinely audited for accuracy because they are being used for pain. And since we're talking about a high-value procedure it's very unlikely that any major inconsistencies or errors would go unnoticed.

MS. TIGHE: This is Lindsey. I'm going to jump in and just circle us back to the reliability discussion because I think we've jumped well into validity at this point. Do we have anything else to say about the precision of the specifications or the reliability testing that was supplied? I don't think we've touched on the reliability testing at this point.

MR. BURTON: So, I'll cover that briefly here. The signal-to-noise analysis that yielded 0.99 for the ACO group and the drug plan group. There was like we had mentioned before an issue with the physician group that only 13 percent of those had sample

sizes large enough to generate reliability.

Just going with the 0.99 that is high reliability but only for those, the larger groups.

DR. GEORGE: Any discussion on that? Liz?

MS. DELONG: Yes. Could you -- I have no idea. I've seen that before and maybe it's my own ignorance, but what did you do for a signal-to-noise reliability test?

MR. CAMPBELL: Sure. So, the signal-to-noise ratio is calculated as a variance of the between measured entities which is considered the signal and the variance within a measured entity which is considered the noise. And then the reliability is estimated using data --

MS. DELONG: So when you say within and between, can you be more specific?

MR. CAMPBELL: Yes. So, it would be like if we were talking about an ACO or a physician group you would look at the within

variance. So within that group statistically is there more noise that sort of drowns out the signal of being able to make comparisons between physician groups.

So, if you can't discern -- I guess the best way to say it, if there's more variability internally within a physician group than there is externally compared to the peers then generally your reliability will be poor.

And so as the reliability approaches 0.7 we can begin to distinguish statistically significant differences between providers from the mean as it approaches 1.

MS. DELONG: Okay, so you're basically looking at the inter-class correlation and -- but you're assuming that you don't have misclassification, right? That you have valid data to work with.

MR. CAMPBELL: That's correct.

DR. GEORGE: Linda?

MS. BRIGGS: So, I just wanted to

echo what we said earlier in that if we're looking at this at reporting the physician group level data the report from the authors of this measure says that only 13.3 percent of the physician groups have an adequate number of patients for reliable measurement. So that's not a very large number of physician groups.

MR. CAMPBELL: This is Kyle again. Go ahead.

DR. GEORGE: Go ahead on the phone.

MR. CAMPBELL: Sure. So, basically the way we do that across the measures is we look to see if there's some minimum denominator or threshold size. Because this signal-to-noise ratio is sensitive to sample size.

So, with that minimum denominator of about 3,650 days or 10 patients within the denominator we do get reliable scores for physician groups. And so that threshold would

have to be considered if the measure were to be used at the physician group level.

DR. GEORGE: Are we ready to vote on reliability? Okay, we'll go ahead.

MS. LUONG: The timer starts now. One for high, two for moderate, three for low and four for insufficient.

Ten voted moderate, ten voted for low and two for insufficient.

DR. GEORGE: We are going to move forward and finish this measure before we go onto public comment. Validity.

MR. BURTON: So with validity we spoke a little bit before we got into this section.

Just as far as the validity testing there was a face validity that was assessed by a technical expert panel in which 80 percent strongly agreed or agreed that the measure was valid.

And given that number and the fact that only face validity was used I think our

highest rating could be moderate. And that the results did demonstrate that this measure is a reflection of quality of care.

I didn't really have too much else on validity. We talked a lot about the data as far as the codes and exclusions. This measure is not risk-adjusted as a process measure. But I'll leave it to the rest of the group for discussion in the purpose of time.

DR. WINKLER: Just sort of pertinent to your previous discussion, this is the point where you want to determine whether the specifications are consistent with the evidence.

DR. HOLLANDER: I sort of said my piece before. I'm not sure this is really adherence. And that speaks to validity.

And although the expert panel that they employed thought it did and I guess NQF has used measures like this before I still don't feel it actually speaks to whether the patient is taking the medications.

And the issues raised by George and Henry, having them out twice means they should be taking it for a year and how it works. So I have major issues with the validity that I don't think I could get at.

DR. GEORGE: And thank you for reminding us of that prior discussion. Any other discussion on it?

DR. TING: This is actually for Kyle. Many people have been critical and made comments, but this is an incredibly important area which is adherence. So if this measure was statins at one year would we know adherence is somewhere around 60 or 70 percent? None of us would have any reliability or validity issues if we could measure adherence to statins at one year.

And this issue of using dual antiplatelet therapy at one year after the stenting is an issue. We know that upwards of 15 to 20 percent stop at six months. And it's correlated with mortality. But the comments

that have been brought up still stand, that there may be some issues with this measure but it's an incredibly important issue in terms of quality of care.

DR. GEORGE: Liz, did you have a comment? Yes.

MR. MARRS: I guess I have just an add-on. The validity issue with the PDC and measuring adherence this way. It is a very validated surrogate marker for adherence. It's used across lots of different disciplines.

And so even though it's not a perfect measure of adherence and it doesn't necessarily make sure that the patient took it it has been validated in lots of other disease states and pharmaceutical studies looking at whether people are adherent or not.

DR. HOLLANDER: So with that in mind I could see there's certain people that the adherence or whether they got their medication should be attributed to. So if

it's at the ACO or the health system level and they're the person who decides the manner in which the patients can get the medication. And I'll go back to is it via mail or do they have to go get it. Then I can see some responsibility. I still wouldn't call it adherence. I'd call it getting the medications or something else.

But if it's a physician group and they're taking care of someone and they have no say over what insurance or how those medications come to that patient I have a real issue with that physician group being responsible for this measure or even be reported with them because they really have no control.

If they're prescribing the best medication that has a class 1A recommendation and it costs too much for a patient making $10,000 a year that patient may take it for two months and stop taking it. And you can't blame the physician group for that. They

don't have a lot of alternatives. And so I have a problem with it at that level.

I don't really have a problem with it at the ACO or the health system level.

DR. KOTTKE: Tom here. I think just jumping ahead for a moment that this would be extremely burdensome for physician groups because they just don't have -- they don't have in their database who fills and who doesn't.

I think for health plans it's quite easy and it's very appropriate. And health plans could do something like hey, you're five months out, you may be thinking about quitting your dual platelets, don't. You know, that kind of stuff.

But my major issue with validity is what Henry and George brought up is that if I'm going along and at four months I think I got by with this old guy and he hasn't bled yet, I'm going to stop his, you know, I'm going to go back to just an 81 of aspirin

there's no way to detect that.

And that well made clinicians very upset. Even if you say well, there's a certain proportion where you misclassify. But clinicians don't like to be misclassified with crude measures.

DR. GEORGE: Are we ready for a vote? I'm sorry.

MS. BRIGGS: So, I would agree that the method might be appropriate and might be used for other measures that NQF does.

However, I think that we have a little bit of a special case here in that we're trying to measure the DES and the bare metal by the same standard. And this is different than saying did you take your statin and other medications like diabetic medications and so forth that may not have a criteria that would be 1 month versus 12 months.

Whereas you want those people to take it chronically. So I think that we have

to take that into consideration.

And again, because those are all lumped together we need to decide whether we need to ask for stratification as a criteria or just not take the measure at this point.

DR. GEORGE: Any other final comments before we vote on the validity? If not we'll vote.

MS. LUONG: The timer starts now. One is for high, two is for moderate, three is for low and four is for insufficient.

Six voted moderate, eleven voted low and five voted insufficient.

MR. BURTON: Feasibility?

MS. TIGHE: Sorry, I'll jump in. The measure was not recommended because it failed to meet the validity criteria. So we'll stop discussion of that measure.

And actually, given the time on the agenda we're running a bit behind. So if we could take this opportunity to see if there are any public comments from those on the

phone. Operator, if you would check and anyone in the room.

OPERATOR: To make a comment please press * then the number 1. No, no public comments at this time.

MS. TIGHE: Okay. It appears there are none in the room so we are -- yes.

DR. PHILIPPIDES: Despite the fact that I brought up several of the issues here that I felt might have torpedoed this I did want to actually -- and I wish that I had said what Peter said.

Which is I do think this issue of taking medications is a huge issue. And I actually don't think that none of it should be laid at the level of the office. Because almost every cardiologist sees a patient after they've had an MI, a stent, within a few -- we try to do it within 8 to 10 days and then again in a few months.

And at that time if you do nothing else you want to make sure that they know what

medicines they should be taking and you get them the medicines. And that means working in conjunction with the ACO and the healthcare system.

So I actually think as do you this is an incredibly important area, not just for aspirin -- you know, this is the beginning of it.

And if they made the tweaks in regards to the bare-metal stents I personally would be much more enthusiastic were it come by our desk again. I don't think we should lose the general concept because of that one detail. I think that would be a loss for us, a disservice to our patients. So I just wanted to echo what you said, Peter.

MR. BURTON: This is Jeff. I'll second that.

MS. BRIGGS: I would agree it's a very important topic.

MS. TIGHE: Certainly the developer has heard that and our staff will

work with him on making these refinements and potentially bringing it back to the committee for review at a later date.

That said we are overdue for a break. I'm looking to the chairs. Do we want to take the full 15 or can we shorten it to 10?

DR. KOTTKE: We can try 10 but it will probably mean 15.

(Whereupon, the foregoing matter went off the record at 3:24 p.m. and went back on the record at 3:41 p.m.)

DR. NALLAMOTHU: So, we're ready to start. So, I'm going to be brief. I'm sure this is going to start up a lot of discussion. So, I'll save my comments for later after listening to your guys' reaction.

But essentially this is a measure related to comprehensive documentation of the indication for PCI among all adults undergoing this procedure. It's a process measure and it's performed at the facility level.

And you can see the text around it is essentially focused on five aspects of a procedure and how well those aspects are documented within the procedural record.

DR. KOTTKE: Linda?

MS. BRIGGS: Okay. So, as they have said there are five different criteria. So it's a component process measure.

In terms of the evidence to go with that they used guidelines as the evidence. The one guideline is the appropriate use criteria guideline from the American College of Cardiology. And that appropriate use criteria guideline was generated by looking at about 180 scenarios that were developed originally to say what would be circumstances under which people would have PCIs. And then an expert panel was convened to judge the appropriateness of use for those particular scenarios.

In order to meet those scenarios they have to use these criteria basically.

So, for patients that had acute coronary syndrome they don't have to meet quite as many of the criteria because they meet it under the acute coronary syndrome and that's reflected actually later on when they looked at the data for this.

But the other patients have to have things, the other items such as the stress tests and the presence and severity of anginal symptoms. And the big one being the stress test.

The other guideline has much more evidence to back that in terms of randomized controlled trials and that is the PCI guidelines from 2011.

So, based on the information that was given about the evidence for this measure the measure does not actually reflect something going on with the patient per se. It's only documentation that we're looking at.

And the assumption is that documentation then mirrors what actually is

done for the patient, and that this would then facilitate quality of care.

There was no quality statement at all for the information that was given in the guidelines. However, at least one of the recommendations that's used is a class 1 recommendation with grade A evidence which would make it multiple randomized controlled trials.

Based on the majority of the information I would say the evidence is moderate for this particular measure.

DR. KOTTKE: Jeff, do you have any comments you'd like to add?

MR. BURTON: Sorry, I was on mute there. No, I don't.

I guess my concern -- I'm not too versed in this area -- is I guess is a lack of making the connection to the outcomes. And if there was -- if there is other evidence that points to how that happens is that -- is that just not available through the guidelines? Or

is there something else out there?

DR. NALLAMOTHU: So, I think that's what's been stated up to this point has been fairly accurate. This is a measure that's focused on documentation.

I think the natural question is how does that relate to outcomes. It's a difficult question because the real focus of this measure is to even get to the point where subsequent measurement can be done. So, it's challenging.

I can tell you that, you know, we did look at individuals where within the criteria -- I'm going to pause here because I want to make sure I explain this in the correct way.

But if you do measure appropriateness which is part of the goal of this measure is to comprehensively document so that can be done, there's really no correlation between appropriateness and outcomes in general. There's very little

evidence.

And that's because appropriateness has very little to do with what we would consider traditional outcomes measurements if you're looking at the basic ones of mortality and procedural complications.

Whether or not that procedure was right for that patient at that time is much more challenging to assess. And so I think that that's been a great challenge for thinking about the link between this and what I would consider traditional outcomes. I hope I didn't confuse everybody.

MS. MITCHELL: Was there a translation of the AUC criteria in two measures? Is this an attempt to do that, or is this completely separate?

DR. NALLAMOTHU: So, to step back. That's exactly -- I mean, that's a great way of putting it.

So, this is essentially a measure

that has developed mainly because of the limitations of measuring AUC. So, it turns out that about one in five, maybe a little less than one in five, all PCIs can even be mapped to AUC.

And then when you look at the elective ones it's much more. It's like about one-third can even be mapped to AUC because the data are just not recorded.

And fundamentally, I mean I think this measure is so important mainly because it moves the field forward with being able to actually even start to assess this really important aspect of care.

Right now these procedures are essentially invisible and we don't have the ability to kind of assess quality in any way.

DR. KOTTKE: Other comments?

DR. AL-KHATIB: I completely agree with that. I actually see a lot of value in this performance measure.

And in fact, if you look at the

Affordable Care Act among many of the quality improvement initiatives that were mentioned is ensuring appropriateness of cardiovascular care is what was mentioned in the Affordable Care Act. I truly see this as a very helpful measure. Hopefully we'll be able to make sure that all the other aspects of it are fine. But I certainly can see a lot of value in this measure.

DR. VIDOVICH: I would just like to echo this. I think it's a very valuable measure.

My question to the developer is how granular will be the measure, the requirement for granularity? What -- will you require that some specific categories are filled in, or anything goes? You mentioned FFR or IVUS for indication criteria. So these synchronize with the AUC at some degree.

DR. NALLAMOTHU: So that's a great question. So, again, there is granularity. The measure itself does get into the specifics

of how that's described.

But to give you a sense it not only requires, for example, the presence of a non-invasive stress test or an FFR, an IVUS, but also in some kind of quantitative terms the results as well.

I think one of the biggest problems has been in some cases, for example, with stress tests there might be documentation that a stress test was performed. But then it's remarkable how that never -- the result of that never actually makes its way into I believe the most important document related to a procedure.

DR. KOTTKE: Any other discussion on evidence? Are we ready to vote?

MS. LUONG: So the timer starts now. One is for high, two is for moderate, three is for low, four is for insufficient evidence with exception and five is for insufficient evidence.

So for evidence 4 voted high, 17

voted moderate and 1 voted low.

DR. KOTTKE: So, we move on. Opportunity for improvement. Jeff?

MS. BRIGGS: Actually, it's me. So just to back up a second. The database for this just to be clear is the CathPCI registry again. So, this is a really large, very rich database that we're dealing with. And we've already discussed how reliable and how it's being used.

Opportunity for improvement. In 2011 they reported that the mean unmappable which means they couldn't find any of those 180 scenarios that based on the amount of documentation that they had for the patient that they were able to map it to one of those scenarios. The mean was 42 percent with the median being 39.5 percent. So, there's a lot of opportunity for improvement.

In 2012 it was slightly better. The lower number actually, the better in terms of the unmappables here. So we're still at

over one-third of patients being unmappable at 37 percent as the mean in 2012 and the median being 35 percent unmappable based on missing data at that point in time.

So there is substantial variation among the various practices that were reporting and the hospitals reporting. They ranged from zero basically to 100 percent. So there was a great deal of opportunity for improvement.

DR. KOTTKE: Jeffrey, any comments?

MR. BURTON: Yes, I just wanted to maybe get a better understanding. I know there was an issue on a prior call about missing data versus other data that was never collected because either a test wasn't done or whatnot. It was a process of care that was broken down.

So, is there any detail that the developer can provide that shows the breakdown of what is actually data that is out there but

the hospital was unable to get due to the fact that maybe there was a stress test that was done somewhere else versus a process of care not being in place to generate the data.

DR. NALLAMOTHU: That's an important gap and that was something that was mentioned in the call as was mentioned.

I think what we've philosophically kind of felt about that is even if the stress test was done let's say by the referring cardiologist at their own office and then the patient ended up going for a PCI that somewhere within that PCI record that stress test needed to be documented. So that's kind of how we would approach that question philosophically.

But we just don't have the ability to kind of tease out how much of this is a lack of results being communicated or the test was never done.

MR. BURTON: Yes, and I'm just trying to get an understanding. I think that

it's valuable regardless of whether or not the data wasn't there for one reason versus another that -- the fact that the data is there during the time of the PCI is the most important part. So, I didn't want to devalue that.

MR. CHIU: And if I can add just one thing to Dr. Nallamothu's point.

I think this measure is a little different than other ones in that there are no exclusions. So in terms of gaming it's kind of a slightly different answer but just to add. There's no gaming.

And it's really simple in terms of what you do with missing data. If there's missing data you basically have failed, you've failed. Because the thought to Dr. Nallamothu's point, you really should be documenting these indications in the long description. Those five points there.

I just wanted to add the missing values should actually be included in the

denominator but you'd actually fail the measure in the numerator.

DR. KOTTKE: So sort of as perhaps an amicus comment that you don't really need the stress test in the record. You need a report or something that indicates this patient had a positive stress test at two minutes with angina. So I'm doing an angiogram.

DR. NALLAMOTHU: Absolutely. It's not the original record but the fact that there was some -- and a lot of times, you know, we, again as a proceduralist myself we make the assumption that, yes, I know it, it's in my brain and I know what I'm doing. But that documentation it turns out is just -- I mean, it's -- as people have mentioned, the opportunities here are pretty tremendous.

DR. KOTTKE: So, any further discussion? Yes, sir.

DR. CLEVELAND: I just wanted to

ask, and maybe Jensen can weigh in too. I know we've struggled with this too in looking at appropriateness and trying to actually data map elements. Are there any plans within the NCDR to data map? Because that would certainly add more robustness to the appropriate use criteria. I mean almost a module type of thing. Do you know? Except that might then take the missing argument pretty well.

MR. CHIU: I think the challenge obviously is -- this doesn't just pertain to just this measure but other kind of measures in NCDR.

So, this one, you know, going through the test and everything once this is endorsed we put it in the registry, in the Cath.

But the challenge always is there are going to be missing data regardless. That is, how much missing data. Unfortunately at this juncture it's a little hard to tell how

much the missing data there is. It's a little challenging to really know all the -- how much truly is missing, how much you can really quantify, because you really don't know what you don't know.

This is a challenge I know STS also has struggled with as well. So it's kind of a challenge.

But the one thing I would add though too just to recall. I don't know if it's in the application, but all the measures that become in Cath and other NCDR registries, the suite, there's a data quality report so that you can have too much missing data. So I don't know the core elements off the top of my head but I'm sure some of these elements are core.

And really that just means that if you have more than a certain percentage that are not being captured you actually are failing. You actually don't get a report. Your site doesn't get a score back to them.

So that's -- we can kind of go back and take a look at what the elements are and then report back on that.

But all registries, Cath probably being -- you know, I don't want to jump the gun but I think probably being more robust than some of the other registries we have. But there is a data quality report that every year is audited. Certain variables.

But the missing again, if a site -- some site or something has too much missing data they don't get a report out.

DR. KOTTKE: Liz?

MS. DELONG: So you have if performed you need the information. But suppose it is performed at an external lab. It was performed but if it's missing you ding the hospital who performs the PCI? You don't know if it was missing unless you link, right?

DR. NALLAMOTHU: So I think the way to think about it is more simpler. Look, whether it was performed or not there has to

be documentation. So if you didn't do it that's not missing.

But what happens is if you didn't do it and you get a PCI and you're asymptomatic and it was just because there was a lesion there then at least you can say that that was inappropriate.

Right now if you don't even have that there that patient falls out. So, again, two scenarios. Somebody who's asymptomatic. Let's say they're not on any medical therapy and they have a limited coronary lesion. That person gets a stent. If they actually record it, if they went to the step of saying, you know what? We didn't even do a stress test that patient gets identified as inappropriate. That patient is at least visible.

What this measure is trying to do is deal with the other side of it which is the invisible. We're in Washington, D.C. so it would be Donald Rumsfeld's unknown unknowns. It's the idea that, you know, if you just

don't even say well, I didn't even record whether it was done or not that person is invisible to the measure as it currently stands. Does that make sense?

MS. DELONG: So, it's actually two items for each thing then. Was it done and what are the results.

DR. NALLAMOTHU: You need to have the results as well too because in some cases like, you know, again a stress test and then not knowing the results of the stress test makes it unmappable as well.

DR. KOTTKE: Any further discussion? We're ready to vote. We're voting on opportunity for improvement.

MS. LUONG: So the timer starts now. One is for high, two is for moderate, three is for low and four is for insufficient. Eighteen voted high, two voted for moderate. Four, sorry.

DR. KOTTKE: Priority.

MS. BRIGGS: Okay. So, as has

been pointed out there are a fairly high number of patients who received in particular elective procedures that are deemed actually inappropriate from one of the studies quoted by the authors that the measure one in eight elective procedures is actually an inappropriate procedure.

And there's a 1.2 percent mortality rate associated with any PCI. So we are exposing patients needlessly in some cases to the procedure if it's inappropriate.

And it's also a fairly costly procedure. In the estimates provided in other documentation by ACC a cath or PCI can cost somewhere about $72,000 by the time you add in the hospitalization component of it. So we are talking about high cost and a possible for harm for patients. So it is a high priority indicator.

DR. KOTTKE: Jeffrey, anything to add?

MR. BURTON: No, completely agree.

DR. KOTTKE: Any discussion? Liz?

MS. DELONG: According to the data that you collected did you see a difference between the inappropriate mortality rate and the appropriate mortality rate?

DR. NALLAMOTHU: So, that's a little bit different. Again, those are people who could even be mapped.

But I do want to emphasize that point about why this is so critical. And using the traditional measures of mortality is probably inadequate.

So, when we've in the past looked within NCDR and we've just mapped based on appropriate indeterminate or appropriate. So all these people could be mapped. And then we just correlated it with simple kind of in-hospital outcomes, typical ones. There's actually very little correlation.

And the way that we interpret that is, and the clinicians here, I mean it would be almost intuitive is that it actually turns

out that it's pretty safe to put in a stent in someone who doesn't need one.

And there's two aspects of care that are being assessed here. And that's why, you know, again, I'm kind of curious to see how this discussion flows. But I really do think that this is such an important first step. Because otherwise it's impossible to assess this other side of it.

DR. KOTTKE: Maybe I can call the question. I think we all believe that putting a patient at any risk whatsoever for no justifiable reason is wrong. So let's vote.

(Laughter)

MS. LUONG: Voting starts now. One is for high, two is for moderate, three is for low and four is for insufficient.

If we all can just point your fob back to me and vote for your number. Yes, thank you. Nineteen voted high and three voted for moderate.

DR. KOTTKE: Scientific

acceptability specifications.

MS. BRIGGS: Okay, so as I mentioned this is a component or a composite measure. So the numerator statement has to do with having all of these criteria in order to be mappable. So there has to be a priority rating, there has to be presence of the documentation of the severity of angina, use of anti-anginal agents, the presence and results of non-invasive stress testing or the fractional flow reserve or IVUS therapy, or estimation. And the significance of the angiographic findings as well. So that's the numerator statement. And if there's a no on any of those then they're not met in terms of having adequate documentation.

The denominator is all patients age 18 and older for whom PCI was performed. There are no exclusions. And in terms of the acceptability for that I think it's reasonable.

DR. KOTTKE: Jeffrey, any

comments?

MR. BURTON: No comments.

DR. KOTTKE: Any discussion? Seeing no discussion let's vote. Oh, I'm sorry. Reliability testing.

MS. BRIGGS: So, in terms of reliability the testing done was signal-to-noise. And with greater than or equal to 80 percent or 0.80 being very good the authors indicated that it was moderate across all centers and it was very good in centers that were more high-volume centers. So there's at least moderate reliability across all centers reporting. And there were over 1,100 centers involved in the data set.

DR. KOTTKE: Jeffrey, any additional comment?

MR. BURTON: I just had one question about the minimum number of cases in a hospital. Was it 10 cases that was used as a minimum to include a hospital in the testing? Or is that -- that seems low to me.

DR. NALLAMOTHU: I think your point's well taken. It is low when we set the standard. Most hospitals were much higher than that.

The issue with CathPCI is that there is at times these hospitals that report kind of in and out. And I have to double-check on this. I apologize, but I'm not sure if it was greater than 10 per quarter as well. Because this reliability testing was done across that. So I think that was the criteria.

But we should know and we should double-check. I'm not sure if Lara or anyone else can check.

DR. KOTTKE: Any further comment? Seeing no action, let's vote.

MS. LUONG: The voting starts now. One is for high, two is for moderate, three is for low and four is for insufficient.

For reliability 7 voted high and 15 voted moderate.

DR. KOTTKE: Validity.

MS. BRIGGS: Okay, in terms of validity there -- the indicators themselves align very well with the data set. So you can actually map across the different indicators that have been used as part of the composite. So that part was very high.

As I mentioned there were a large number of sites involved in the testing for validity. There were in 2011 1,146 sites and in 2012 the data they presented was from 1,178 sites. So, there's a great deal of patients involved.

In terms of potential threats to the validity there is a degree of threat in the sense that there was the all-or-nothing failure to meet the measure has to do a lot with missing data related to the stress testing. And in some cases it was almost 40 percent of stress test data missing.

And part of the criteria there if you go back to the actual PCI registry itself

and look at the data entry points, it can be a stress test or IVUS report from up to six months before.

So, there's maybe some mechanistic kinds of problems with entering that data. Again, the cath-ing interventionalist may well have that data in his head, but if it's not entered into the PCI registry, if the data never gets there from whatever center did the particular stress test then it's recorded as not met and not documented. So it's then not meeting the criteria. And so something probably needs to be looked at to address that particular issue.

DR. KOTTKE: Jeffrey?

MR. BURTON: No comment here.

DR. KOTTKE: Anybody else? Comments? Seeing no -- oh.

DR. WINKLER: I have a question. And maybe it's just I'm missing something.

This to me seems more than just documentation. So, I want to be sure I

understand what the measure result is.

Linda, you said that in 2011 the mean result of unmappable patients was 42 percent. So, you know, the performance on the measure was 58 percent.

But we've got 40 percent of people that are unmappable. And are they unmappable just because they didn't document? Or it's possible that they're unmappable because they don't meet the criteria, the appropriateness criteria. I mean, is it purely documentation, or are you capturing both together? Those that are inappropriate as well as those that are sloppy in their documentation. Is this picking up both of those?

DR. NALLAMOTHU: The best way I can kind of point this out is so the -- probably the most well known paper associated with this is a paper by a good friend of mine, Paul Chan and his colleagues in JAMA. And I think it was around 2010 or so.

But I'm just going to read from

here the figure. So when PCIs are excluded it turns out that that's -- out of this there were 600,000 PCIs that were done. One hundred thousand of them had to be excluded because they couldn't be mapped to the AUC.

About 50,000 of those were non-acutes with no prior stress test. And of those because of that about half of those were unable to be matched to the appropriate use criteria specifically because they didn't have a prior stress test.

About 40,000 of them had a prior stress test documented but there was no ischemia risk specified, making it difficult to assess what the actual value of the procedure was.

So, you know, a lot of this is tied to the stress test, no question about it. That's the documentation that's probably the most challenging and difficult to overcome here.

But I think it is interesting

because the -- it gets at kind of what you're mentioning, whether again, without that information it's just very difficult to use the AUC.

So, there would be about 10,000 left where it was because of other reasons, either other missing data elements or the fact that it was one of these -- you know, I mean they looked at about 200 scenarios and I guess there are other scenarios besides those 200. But for the most part they're of small proportion.

DR. WINKLER: And the reason I raise it is because the word "documentation" is going to raise a red flag for certain stakeholders who feel that documentation measures are pretty minimal if you will. You know, did you document symptoms. Did you document this or document that. And I wonder if that in the title is maybe misleading, that there's more to this measure than simply documentation, that actually we've got a lot

more appropriateness built into this measure than whether they check the box or not? And that's -- I'm just wondering if this is going to get perceived with that in the title as is it just another documentation measure as opposed to something quite a bit more robust.

DR. RUGGIERO: The question I had is if you don't have an objective study maybe percent lesion as written in a chart would be a documented -- not necessarily a documented failure given the story. So, I think your point is well taken.

DR. KOTTKE: Leslie?

DR. CHO: I think it's a very, very important measure for many reasons. I think, number one, it's the amount of PCIs done in this country without really appropriateness.

And I think that NQF, one of the roles of NQF is really to guide clinicians into appropriate criteria. More than just did you get an aspirin, did you not get an

aspirin, did you get a statin or not.

But I think the missing data component is appropriateness criteria came out in 2009. And it's been a moving target. And many of the hospitals are just figuring out how to put these things into a database. And that's why there's some missing variables.

For example, FFR is not included in the 2009 appropriateness criteria. And so I don't think this measure is diminished because of the missing variables.

DR. KOTTKE: Judd?

DR. HOLLANDER: Trying to be forward-looking on this there's another set of appropriateness criteria that's coming out now for low-risk chest pain, and coronary CTA is prominent in that. It doesn't show up as even something that's being collected here.

And although one could argue if you have an 80 or 90 percent lesion should you go to cath next it's certainly happening. And so I would urge you to at least collect that

data and record that as well because it's getting more commonly used these days.

DR. NALLAMOTHU: The only thing I would say to that is that I think that these are all people who ultimately are going to have a PCI and they have an angiographic, you know, it's an invasive angiographic as opposed to a coronary CTA.

And I think the question is really you should probably still get a functional assessment in somebody who. You know, because you're absolutely right.

I think, you know, if you're thinking about a documentation measure of whether they should even get a diagnostic cath coronary CTA should be right up there with a stress test. Does that make sense?

DR. HOLLANDER: Well taken. Once you have the diagnostic cath you go by that.

DR. GEORGE: I would just add that I think oftentimes the documentation of appropriate use is so important. Without it

you're not able to develop appropriate outcome measures.

DR. KOTTKE: Further discussion or are we ready to vote? Looks like we're ready to vote on validity.

MS. LUONG: Voting starts now. One is for high, two is for moderate, three is for low and four is for insufficient.

Six voted for high, 14 voted for moderate, 1 voted for low and 1 voted for insufficient.

DR. KOTTKE: Feasibility.

MS. BRIGGS: So, this is a new measure so it has not been used by itself at this point in time. There is another NCDR indicator, the 30-day mortality that's being tested apparently presently. And so it was felt that that would be a good surrogate for the testing, for this related to the PCI data registry.

I think given that it is the PCI registry and that we're using that for a

number of other indicators that it is a feasible study.

DR. KOTTKE: Jeffrey?

MR. BURTON: Nothing much other than the fact that it's the CathPCI registry again. You know, we have a large majority of hospitals participating but some that do not which would give them access to the data. But that's been mentioned before.

DR. KOTTKE: Henry?

DR. TING: Yes, so I've reserved my comments for the feasibility section, not the reliability and validity section.

But just for me to understand this, Brahmajee and Jensen. This is about improving documentation of these criteria so you can map more procedures to appropriate, inappropriate, or indeterminate. It's not really actually a measure of how many procedures that we're doing are actually appropriate, it's just mapping the ones that we can't map right now to appropriateness.

And the reason I ask that is it's almost -- I don't know if this is the first time NQF is evaluating a measure like this for quality and performance. Because appropriate use criteria are based almost on opinions of 16 to 20 experts in a room, 180 clinical scenarios using a RAND modified Delphi technique where you vote and you don't even discuss the case. So it's very much expert consensus.

And you wonder why any payer would actually pay for a procedure where there's no indication of why it was done, you know, be it a CT scan or a PCI. And whether this is an NQF performance or quality measure as opposed to why are we paying for this if there's no documentation that the person needed a procedure.

Which gets back to Tom's statement. You know, if you don't need a procedure you shouldn't be exposed to any risk. So, I'm just asking that question under

feasibility how does that fit within the NQF measures.

DR. HOLLANDER: I was following up on comments from before, say, that maybe this isn't about documentation. Maybe this is about the ability to determine appropriateness. Right? Because that's what it's all about.

And then I would say that that probably does fall within NQF if the title is changed to reflect that.

DR. WINKLER: In terms of NQF we would be totally delighted to have measures of appropriateness.

I agree that we may have to flex a little bit of the criteria because it isn't the traditional structure-process-outcome sort of thing. And you're right on expert consensus.

Under the current thing if you were talking about evidence this would be one of the very best reasons for an exception.

But we probably -- if indeed hopefully this is the beginning of a new type of measure. We'll have to adjust the criteria to account for it.

But by no means -- our stakeholders would be delighted to have an appropriate use criteria measure. No doubt about it.

DR. TING: Again, I'm not trying to develop a new measure, but why wasn't something just like percent of procedures that are deemed appropriate the measure? As opposed to trying to get the ones that are unmappable, Brahmajee.

DR. NALLAMOTHU: Well, I would say that ultimately I think that that's an important kind of goal to shoot for.

But when you have one-third of the PCIs and sometimes at some centers 100 percent of the PCIs unmappable I think it really -- it sets a disincentive for being able to -- I mean the easiest way to meet criteria is just don't say it.

And I think what's more interesting, and maybe I'm misunderstanding, but I do think what I hear about what you're saying is is this even a quality measure or is this just like a standard for getting paid. And that's a broader question.

I do think it is within the purvey of the NQF, but that's a personal opinion.

DR. TING: So for -- and just, again, I don't want to sort of say anymore. It's my last comment.

New York State, for example, if a procedure is not deemed -- if it's inappropriate and SNAP as such it's actually not reimbursed if you're Medicaid in New York State. I mean that's already been a payer decision state level.

MS. SLATTERY: So we're talking to New York State about that and the appropriate application of our appropriate use criteria or potential inappropriate.

I do think that that's an

important distinction though. The appropriate use criteria do not function like traditional performance measures. And one of the discussions that happened earlier with our first measure was what's the target. And when it's a performance measure we know and think it's fairly reasonable usually that they can get to 100 percent. That is not the case with the appropriate use criteria.

More to the point, if we were to even attempt to put forward any type of performance measures with a target around appropriateness we would need to understand and have more complete reporting going on with patients to say well, what really is the target that we think we could reasonably move the hospitals towards.

Which means better documentation, ergo why we're putting this measure forward. Because we need better documentation from the hospitals. Just reporting out appropriate use criteria is not sufficient to get them moving.

DR. NALLAMOTHU: And just to build on what Lara said is imagine if we came to you -- we do have a measure that's like that. But imagine if we came to this group with that measure and we said oh yes, by the way, about one-third of them we can't even tell. I mean, that would definitely be an uncomfortable discussion. So I think that this is that first step.

DR. KOTTKE: Ready to vote on feasibility?

DR. AL-KHATIB: I just wanted to add one quick comment, that I completely agree and I completely see this as part of the quality improvement initiative here.

Because even if it's just documentation you're getting the healthcare providers to think about these things. And to question do I have an indication here. So I certainly see it fitting into the quality improvement initiative.

DR. KOTTKE: You could define

documentation as part of the process. I mean, it's like washing your hands before you cut somebody open. It's a process.

Okay, ready to vote on feasibility.

MS. LUONG: The timer starts now. One is for high, two is for moderate, three is for low and four is for insufficient. Thirteen voted high and nine voted moderate.

DR. KOTTKE: Usability and use.

MS. BRIGGS: I think I actually already reported on this under the feasibility, but this is not currently being used. They're piloting a surrogate of 30-day risk for readmission. And there's no public reporting of this currently.

The documentation piece I think is again useful information. Again, it's probably a good first step to getting to the actual appropriate use.

DR. KOTTKE: Jeffrey, any comment?

MR. BURTON: No comments here.

DR. KOTTKE: Discussion? Any discussion? Seeing no motion, let's vote.

MS. LUONG: Voting starts now. One is for high, two is for moderate, three is for low and four is for insufficient information.

The usability criteria has 13 for high, 8 for moderate and 1 for insufficient information.

DR. KOTTKE: Final vote, overall suitability.

MS. LUONG: The timer is now. One is yes, two is no.

Twenty-one voted yes, one voted no.

DR. KOTTKE: So you're batting 500. You could play for the Angels.

(Laughter)

DR. NALLAMOTHU: Thank you.

DR. GEORGE: Next we are moving onto measure 2459 in-hospital risk-adjusted rate of bleeding events.

MS. TIGHE: Do we have anyone from ACC joining us for this measure?

DR. MASSOUDI: Okay, I know we're behind schedule so I'll give you like three sentences.

But this is a measure that uses again the CathPCI data registry to report risk-adjusted rates of periprocedural bleeding after PCI using a validated model that's been published in JACC Interventions by Rao and colleagues.

This is unlike the previous measures we've been discussing which are process measures and in the last case sort of an appropriateness measure, this is an outcomes measure. Again, using a validated risk-standardized model.

And that's all I'll say unless there's -- okay.

DR. AL-KHATIB: I guess I'll delve into it. So, as was stated unlike all the other measures that we've discussed today this

is an outcome measure. I'll get to the evidence here.

The developer provided evidence, or at least results from several large studies to make the case that there are processes of care that can influence the outcome. So, they mentioned a study that was published by the group at the Mayo Clinic that determined that there were certain factors related to the sheath size, intensity, duration of anticoagulation with heparin and procedure time that are independent predictors of complications. Talking about several other studies as well highlighting really several factors that are linked with increased risk of bleeding. And certain things that we certainly could do to try to minimize the risk of bleeding. Based on that I think the level of evidence is high.

DR. WINKLER: Just to remind everybody that what we're expecting for evidence for outcome measures is not the same

as for process measures. And simply are there things can we do to influence the outcome. It's a straight yes or no on the evidence.

DR. GEORGE: Any discussion on the evidence? We'll move to a vote then.

MS. LUONG: The timer starts now. One is yes, two is no.

We have 100 percent 21.

DR. AL-KHATIB: So moving onto the opportunity for improvement. They did a study within the CathPCI registry and they certainly demonstrated a gap in care. The risk of bleeding, at least the mean risk of bleeding was 5.5 to 5.6 percent. That may not seem that all impressive. It is significant to me.

Although I don't expect that risk to be zero we really have to strive to be as close to less than 1 percent as possible.

The concern there though is the variation in the risk of bleeding where they clearly said that the distribution of hospitals show that there are some sites with

excellent performance and other sites with rates of bleeding that were 80 percent or greater than expected risk of bleeding. So with this information in mind I think that there is certainly a significant gap in care and a tremendous opportunity for improvement.

With regard to the disparities question also the developer highlighted that there were some statistically significant differences by gender, race, insurance status, but that the absolute rates after patient-level adjustment were clinically marginal except for gender which is a strong risk factor for bleeding. So hopefully this could be reported at least by gender as a performance measure.

DR. GEORGE: Any comments? Performance gap. Hearing none we'll move --

DR. JAMES: And this is more of a question. Because I think of this as being analogous to the CLABSI and the CAUTI types of things. Do we have standards that would

preclude -- that would help push us down to a zero rate of bleeding?

DR. AL-KHATIB: That's what we discussed under evidence in terms of like are there any processes of care that can help us lower that risk.

And as I said they actually cited a lot of papers where several risk factors have been identified that you could base -- knowing about those risk factors you could be extra cautious, extra careful. Talking about like personalized medicine and what have you.

Maybe even question whether that patient -- what kind of anticoagulation you need to give them, things like that to try to go for the medications that are associated with the lowest risk of bleeding and things like that. So certainly there are things that can be done to lower the risk.

I'm not aware of a checklist.

DR. MASSOUDI: Maybe not a checklist but there are tests of approaches

underway where one could personalize the use of bleeding avoidance strategies like the use of bivalirudin closure devices and radial access based on a patient's individualized bleeding risk. So there are sort of approaches in place where that could be integrated into care. That's obviously not the goal of this measure here but that's been tested and performed. And published, yes.

DR. GEORGE: Any other comments on the disparities and gaps in care? If not we'll move to a vote.

MS. LUONG: The timer starts now. One is for high, two is for moderate, three is for low and four is for insufficient.

Nineteen voted high, two voted moderate.

DR. AL-KHATIB: Moving onto priority. Yes, I believe this addresses a significant health problem. As I mentioned when I talked about the initial measure related to PCI, you know, CAD is a very

prevalent condition. PCI is very commonly done associated with high costs. And I believe that this measure fulfills the priority criterion.

DR. GEORGE: Any discussion on priority? If not we'll move to a vote.

MS. LUONG: The timer starts now. One is for high, two is for moderate, three is for low and four is for insufficient.

Seventeen voted high and three voted moderate.

DR. AL-KHATIB: Okay, so moving onto scientific acceptability specifications. So the numerator is all patients 18 years of age and older undergoing PCI and developing post-PCI bleeding.

The definition of bleeding was very specifically provided, bleeding event within 72 hours. And all definitions use a greater than or equal to 3 grams per deciliter drop in hemoglobin or transfusions, or an intervention to stop the bleeding, or

hemorrhagic stroke, or tamponade, or post-PCI transfusion.

And then the exclusions were NCDR registry patients who did not have a PCI obviously. And patients who died on the same day of the procedure. Patients who had CABG during the admission. Patients with pre-procedure hemoglobin of less than 8.

And the denominator were all patients 18 years of age and older undergoing PCI. And as was described by Fred this also uses the CathPCI registry.

I personally think that the construct of the measure is very reasonable. This definition of major bleeding is very much in line with the accepted definitions in the field. I personally don't have any concerns about the specifications, definitions, or coding.

DR. GEORGE: Leslie?

DR. CHO: As a practicing interventionalist one of my pet peeves is this

bleeding criteria. And I just want to, you know.

And my number one thing is that this excludes bypass patients, but it doesn't exclude patients who have had, you know, for example, go onto have TAVR. Go onto have structural, you know, balloon valvuloplasty and things like that.

Because the criteria, I go through this with my NCDR registry nurses all the time. So if I do a PCI and then two days later they go for a balloon valvuloplasty I get dinged on my PCI. Or, if they go for a permanent pacemaker I get dinged on my PCI.

So, there's all these sort of scenarios in which I think it's not a trivial amount of patients only because as we're doing more and more valvuloplasties on older and older patients I think this exclusion criteria -- bypass is good, but I think we need to think about other ones too.

DR. GEORGE: Any comments, Fred?

DR. MASSOUDI: Yes, I mean, I hear what you're saying. And we're doing more and more TAVRs as well. And I'm sure that it will disproportionately influence centers that are doing those sorts of things.

I think that that's good feedback and certainly something that could be accommodated in future iterations of the bleeding model.

DR. CHO: I think if you're doing PCI and then you're going onto have other procedures I think that there should be some amount of leeway for that.

Especially centers like ours at the Cleveland Clinic, or Mayo, or Duke, or wherever. I mean I think those are big issues.

MS. SLATTERY: So, we would agree and that is noted for our version update. It necessitates us updating the data set which we don't do without a lot of pain and trepidation.

Specifically with TAVR I mean that's one of the challenges also when you have a rapidly adopted procedure. How can our registries keep pace and also while we're evolving appropriate use criteria and a whole lot of other things going on.

The other reminder. We do intend this for public reporting. It is a voluntary public reporting program. So for -- that may still be isolated to specific sites. And so there's the chance that they will choose not to voluntarily report that data.

We don't intend that to be a judgment on a hospital. That doesn't mean that we are ignorant to the perception of if a hospital chooses not to it's then left to them to explain why they chose not to which can include their program is at a different place with where the measure is able to reflect the care they're providing.

DR. CHO: I mean, I think it's important for the measure to be accurate only

because insurance companies like Anthem are now going to start their payment based on the bleeding criteria.

And centers, big centers like ours and other centers across the country will be dinged because we do these high-risk procedures and we do combine, piggyback on each other. So I think it is important.

DR. MASSOUDI: Point well taken. Thank you. A risk adjustment for a lot of the characteristics might underlie that. So that may account for some of the variability.

However, at the end of the day the point is well taken that there are procedures like, you know, again an exclusion for bypass surgery is done specifically because the bleeding definition includes blood transfusions. So the point, as I said, point is well taken.

DR. HOLLANDER: I had a process question. Is this a composite outcome since it's a bunch of bleeding from different

sources? Or is bleeding one thing? And so I raise that.

And the reason I raise that is there's one thing -- and the interventionalist can tell me if I'm thinking about this wrong. Like tamponade being in there I think of as a more mechanical problem than a spontaneous bleeding problem. Is that wrong?

DR. AL-KHATIB: Related to the procedure. I mean, so this is bleeding that's related to the procedure. So that's why they've mostly thought about the major bleeding complications that could be related to the procedure.

DR. GEORGE: And I don't think this was intended as a composite.

DR. MASSOUDI: I mean, that's really a technical question that I'd bounce back to the NQF. I mean, ultimately it's one outcome.

DR. WINKLER: Our most recent composite report talks about a type of measure

that's a version of the all-or-none which is any-or-none which we often see with complications.

And so it's a bit of a change and really it's a matter of how do you tag these measures. A measure is a measure, whether you call it a composite or not. So it does have characteristics of it.

ACC would prefer not to call it a composite. We allow them to say no, it's an outcome measure. Fine. So, it's a little fuzzy.

DR. VIDOVICH: I just had a question since we talked about the bleeding. There's a variety of bleeding avoidance strategies, you know, and there's access-related bleeding and non-access related bleeding.

Would it be helpful if you maybe differentiated between those two in this measure? Because radial access may impact the access-related whereas use of bivalirudin may

impact non-access site bleeding. Would that be helpful in reporting and then outcomes?

DR. MASSOUDI: Yes. I mean, it gets to the point of sort of feasibility. You know, trying to identify what one person might consider procedural related bleeding versus non-procedural related bleeding. And so the definition is intended to try and identify with the best sensitivity and specificity possible, acknowledging that there will always be a little misclassification in anything that you do, bleeding that's related to the procedure in one way or other.

DR. VIDOVICH: And this will be in-hospital bleeding, correct?

DR. MASSOUDI: That's correct, yes.

DR. AL-KHATIB: So, we're not voting yet. Let me talk about reliability testing.

I thought the reliability testing was excellent because they performed the

testing at the level of the measure score as well as the data element. And they really provided a lot of details about how they did that. And I have no concerns about the methodology that they used.

They also reminded us of all the quality assurance initiatives that they have within the NCDR program. And so as I said overall I had no concerns about the testing.

Given that the testing was done at the data elements level and the measure score level I would rate this as high.

DR. GEORGE: Any discussion on reliability? All right, we'll vote on reliability.

MS. LUONG: The timer starts now. One is for high, two is for moderate, three is for low and four is for insufficient.

Fifteen voted high and seven voted moderate.

DR. GEORGE: We'll move onto validity.

DR. AL-KHATIB: So compared with the first measure that I presented I think the developer did a better job with the validity here because they actually did some testing. They talked about again the audit of the data and showed how the data elements are valid.

They talked about face validity and described it as content validity of this process. And they really provided a lot of detail on how they did that. I felt that they provided a very good argument for the fact that their data and the data elements are valid. The testing was pretty reasonable and convincing to me. Going through all of this here, making sure that I didn't see any concerns. And I actually had no concerns at all about the validity and I rated it high.

DR. GEORGE: Any discussions on the validity? If not we'll move to a vote.

MS. LUONG: The timer starts now. One for high, two for moderate, three for low and four for insufficient.

Seventeen voted high and five voted moderate.

DR. AL-KHATIB: Moving onto feasibility. I think we've discussed this now several times with regard to using the CathPCI. I think it's pretty feasible and I have no concerns about feasibility.

DR. GEORGE: Any comments on feasibility? If not we'll move to a vote.

MS. LUONG: The timer starts now. One for high, two for moderate, three for low and four for insufficient.

Nineteen voted high and three voted moderate.

DR. AL-KHATIB: And last but not least is usability. So in terms of current use of the measure it's not publicly reported. I think that's what's planned.

It is being used within a program called the Blue Distinction Centers for Cardiac Care. Again, the sponsor is Blue Cross Blue Shield.

And ACC again mentioned their program that they started in July of 2013 where they gave hospitals the opportunity to voluntarily report on some measures.

And although this was not the particular measure that they used they said that they intend to incorporate this measure in their voluntary program.

In terms of unintended consequences the developer mentioned the most vulnerable aspect of this measure pertains to physician transparency and willingness to report and record adverse events.

The one thing that I would add is the potential for physicians to avoid doing PCI procedures on high-risk patients. We did talk about risk adjustment. And although that should alleviate that issue I'm not sure that it would take care of it completely. But I don't see this as a major issue.

DR. GEORGE: Discussion on usability?

DR. TING: So, quick question. This is at the hospital level and not at the clinician level. So there's probably -- vis-a-vis our first conversation there's probably differences within facilities for individual operators. But this is a valid measure and a feasible measure just for the hospital measure though.

DR. GEORGE: Other comments? Judd, did you have a comment? Any other comments? All right, we'll move to a vote on usability.

MS. LUONG: The timer starts now. One for high, two for moderate, three for low and four for insufficient information.

Sixteen voted high, five voted moderate.

DR. GEORGE: All right. We will move to a vote on overall acceptance of this measure.

MS. LUONG: The timer starts now. One is for yes and two is for no.

One hundred percent voted yes, 22.

DR. KOTTKE: Ready for 0133?

DR. MASSOUDI: 0133 is the in-hospital risk-adjusted mortality rate in patients undergoing PCI. This is our oldest -- the NCDR's oldest risk-adjusted outcomes measure. It's actually been endorsed in two previous cycles so this is the second renewal. Yes, it's an alum of the process.

This measure includes all -- is intended to include all adult patients, i.e., older than 18. And applies a widely validated and repetitively validated risk adjustment model to assess in-hospital mortality in all comers.

The distinction between this parenthetically and the upcoming measures which assess 30-day mortality is that 30-day mortality is restricted to those patients for whom claims data for mortality are assessable. So this measure can be calculated irrespective of the availability of subsequent claims data.

DR. KOTTKE: George.

DR. PHILIPPIDES: So, this measure basically allows for benchmarking against national aggregates and against other hospitals with similar PCI volumes as your own hospital.

And it's basically an effort to analyze best practices and disseminate them to try to improve practice.

As Fred mentioned this is derived from the very large and robust CathPCI registry using a big population looking at many variables and after regression sort of paring them down to I think the eight variables that have sort of the most impact on risk of mortality.

This is an outcome measure. And the developers did a very nice job of linking different activities and processes of care to this overall outcome.

And the bottom line here is that by understanding personalized risk of the

patient it allows for personalized care and improvement in the care of that individual patient.

So, with regards to evidence I thought it was pretty strong and no problems with that.

DR. KOTTKE: Mary, any comment additional? Any -- we lost Vy. We have a pinch-hitter here.

MS. MITCHELL: I have a process question. So if this is the third time that this measure has gone and been presented to NQF is there any particular reason why we need to go through every single segment? Was my point. And voting on it.

DR. WINKLER: Simply because that's just a standard maintenance procedure and we don't really want to treat different measures differently.

You're right, the good measures, you know, continue.

DR. PHILIPPIDES: I'll use that as

an excuse to go really fast.

DR. KOTTKE: Right.

(Laughter)

DR. PHILIPPIDES: Thank you for that.

DR. KOTTKE: Can we finish this by 5, George.

DR. PHILIPPIDES: Let's vote.

DR. KOTTKE: So we're up for a vote on evidence.

MS. LUONG: So the timer starts now. One is for yes and two is for no.

I think we're missing -- we're missing a few. If you can just keep pushing real quick. Thank you.

One hundred percent which is 21 voted yes.

DR. KOTTKE: Opportunity for improvement.

DR. PHILIPPIDES: So the developers analyzed a huge database from 2011-2012 with about 1 million patients. And they

found a performance gap, 10th percentile performance 0.7 risk-adjusted mortality. And the 90th percentile was up at 2.7 percent. So I think they correctly identified a room for improvement and opportunity there. So we thought that that was pretty strong.

In regards to disparities there were some statistically significant disparities in regards to race and gender and other populations. But when it was risk-adjusted those became very, very small.

The only thing that did seem to come out and was a little bit more robust -- I think you mentioned this too -- were private insurers and suburban hospitals versus urban hospitals.

In regards to gender and race the differences were very small when risk-adjusted. So again there was no compelling reason to think about stratifying anything and the disparities shouldn't really get in the way here. So that was okay.

DR. KOTTKE: Mary?

DR. WINKLER: Just a point to the developers. Since this measure has been in use for such a long time it would be really interesting to know however far back you have data to see trend over, what, the last decade? I mean, is it really something that everyone should really feel good about, that significant improvements in PCI mortality have really improved over the decade?

It's a great story to tell if we've got a nice downward trend. So, for a longstanding measure like this it's really nice to have if it's available.

DR. KOTTKE: Is there a comment down there? A couple of comments down there?

MS. DELONG: Yes, I want to second that. I think that's important for most of these measures. If there is a trend that we can see it will be helpful to track its utility over time.

DR. KOTTKE: Where are we? Time

to vote for opportunity for improvement.

MS. LUONG: The timer starts now. One is for high, two is for moderate, three is for low and four is for insufficient.

Eleven voted high, eight voted moderate and two voted low.

DR. PHILIPPIDES: Okay, in regards to priority. And my being brief, CAD, MI, PCI - high priority.

(Laughter)

DR. KOTTKE: Sounds good.

DR. PHILIPPIDES: Any questions?

DR. KOTTKE: Mary says nothing. Anybody feel the urge to do anything but vote? So let's vote.

MS. LUONG: The timer starts now. One is for high, two is for moderate, three is for low and four is for insufficient.

DR. KOTTKE: I think you forgot to mention death, George.

DR. PHILIPPIDES: Say again?

DR. KOTTKE: You forgot to mention

death.

MS. LUONG: Twenty voted high and one voted low.

DR. KOTTKE: Acceptability.

DR. PHILIPPIDES: Should I do specifications? The specifications are pretty clear. The numerator statement was, as mentioned, patients 18 or older with a PCI procedure who expired. Denominator are patients 18 years of age or older with a PCI procedure performed during that admission.

There were two exclusions. One, if you got cath but didn't have a PCI. So we're looking at basically patients who had a PCI. And secondly, if you were transferred to another facility on discharge you were excluded. And that's pretty much standard fare.

In regards to reporting on the data source and specifications I think we sort of went over the model. And are we to reliability? Okay.

So, reliability was done here at the data element level and the measure score level. So testing of the performance measure level was conducted with a signal-to-noise analysis.

And it appeared overall the score was good, 0.7 or greater. But when it was broken down by high-volume and low-volume centers it was acceptable for the high-volume centers but not so much for the lower-volume centers. And that was something that I think, Mary, you brought up or somebody did during our discussion. So that requires perhaps Fred addressing it.

In regards to the data element testing that was conducted with a test/retest approach. Basically anybody who was admitted twice within 2012, during that period, and got two PCI procedures were compared to each other. And basically -- were basically classified. And it looked as though misclassification of data elements was very,

very low, less than 3.5 percent across the board. So, actually pretty good.

So, I think the only thing to really talk about in regards to reliability is what to make of the data on the low-volume centers. Any comments?

DR. KOTTKE: Anything more, Mary? Anybody else have comments? Seeing none, let's vote on reliability and scientific acceptability.

MS. LUONG: The timer starts now. One is for high, two is for moderate, three is for low and four is for insufficient.

Eleven voted high and eleven voted moderate.

DR. KOTTKE: Validity, George.

DR. PHILIPPIDES: Okay. No empiric validity testing was conducted. The developers felt that none was necessary other than establishing content validity because the model looking at mortality is of unquestioned importance and is readily assessed.

In regards to content validity the developer did describe the same sort of method that you described, Sana, where they basically looked at data that was coming in and if it didn't have -- if it wasn't complete and also didn't have accurate data as assessed by comparison to the medical record it was given a yellow or a red statement.

Only if it was complete and accurate based on that comparison did it get a green stamp. And they make the point that only sort of green-stamped data packets were allowed into the EDW. So that was their way of looking over this.

I don't know much about the system but it seems like a large number of data packets are checked that way and it's been used for a long time. So it seemed a reasonable way to get a content validity.

But overall there were no numbers attached to that, no sensitivity or specificity. And they relied I think on face

validity if I'm not mistaken.

DR. KOTTKE: Mary, anything?

DR. GEORGE: I'll just add that the missing data was imputed with mean or median.

DR. MASSOUDI: Yes, generally I think that -- I'd have to look back at the model, but the general approach that's been used is that for infrequently missing values missing data are imputed with the median or most common value for categorical values.

Substantially missing data are generally not considered candidates. And for intermediate missingness multi-variable imputation is typically used. Again, I'd have to look back on the specific specifications for the individual variables involved here. But that's the typical approach.

DR. GEORGE: I think there were only two that had any significant missingness.

DR. PHILIPPIDES: Yes, that's right. It was GFR and EF. And they did a

good job of imputing that, things that would make sense clinically.

Exclusions were less than 1 percent. And they even went so far as to derive a C statistic which was really good at 0.93. So I think all of the threats to validity and validity testing were appropriate for this.

DR. KOTTKE: Comments? Hearing none, let's vote.

MS. LUONG: The timer starts now. Voting options include one for high, two for moderate, three for low and four for insufficient.

Eleven voted high and eleven voted moderate.

DR. KOTTKE: Feasibility, George.

DR. PHILIPPIDES: So, we discussed the feasibility of using this registry before. Other than the fact that not all of the elements are always in the EMR no matter what anyone says, and the fact that you have to pay

a small fee to be in the registry, this is a registry that's been used for a long time with good results. And it's got a long track record. So I think that this is feasible.

DR. KOTTKE: Mary? Nada? Let's vote.

MS. LUONG: The timer starts now. One is high, two is moderate, three is low and four is for insufficient.

Eighteen voted high, four for moderate.

DR. KOTTKE: Usability and use. George?

DR. PHILIPPIDES: So, as mentioned before this measure is not being publicly reported. It is being used as a feedback mechanism for hospitals within something called the Blue Distinction program. But I guess there are plans to sort of expand that.

In regards to improvement over time I heard you guys talking here. There are some papers showing that we have improved our

performance over time.

The data that was provided by the developer though didn't really show that. It looked like at least within two cohorts, 2011 and 2012, I think it was roughly the same as far as performance.

Now, it could be that the patients were sicker and that flew under the radar screen, but we didn't see data that showed improvement over at least that one year.

In regards to unintended consequences there were concerns in the past that this risk score did not do an adequate job of assessing risk to high-risk patients.

And that might lead to sort of risk-averse behavior on the parts of interventionalists who basically say look, every time I do a high-risk patient they don't score it high enough and then I get dinged.

But I believe the registry went back and sort of did another analysis and added one or two other risk factors to it. I

think getting into cardiogenic shock maybe, or -- something else might have been added later. And now it appears that this is valid at low risk and high risk when looked at. Do I have that right?

MS. SLATTERY: We did do an exploratory analysis to validate whether that perception was correct or not, breaking it out into different risk groups.

Actually, what we found at the end was the model actually held up fairly well for the in-hospital one. You may be thinking about the pair of models that are about to come up that are harmonized with this measure in terms of breaking it out by shock and cardiogenic shock.

DR. PHILIPPIDES: Okay.

DR. MASSOUDI: But the variables themselves haven't changed. There's actually pretty strong evidence that the model performs well at all, you know, across the spectrum of risk.

DR. PHILIPPIDES: Okay. So no issues there.

DR. KOTTKE: Any discussion? I mean, we all know that our patients are sicker than everybody else's.

DR. PHILIPPIDES: The ones that we intervene on.

DR. KOTTKE: Yes. Let's vote on usability and use.

MS. LUONG: The timer starts now. One is for high, two is for moderate, three is for low and four is for insufficient information.

Nineteen voted high and three for moderate.

DR. KOTTKE: Let's vote on the overall.

MS. LUONG: The timer starts now. One for yes and two for no.

Can we just re-press your votes again? There you go. One hundred percent consensus, 22.

DR. KOTTKE: Thank you, George.

DR. GEORGE: Moving onto the next to the last measure of the day, 0535.

DR. MASSOUDI: Just very quickly. First of all, I think my patients are sicker than yours, Dr. Kottke, but I could be wrong.

So, these last two measures are intended to be used as a pair. I know they'll be discussed separately but it's an important issue to keep in context.

And these are 30-day all-cause risk-adjusted mortality following PCI in two distinct groups of patients. The first being patients -- the first one is going to be 0535 which is patients without STEMI or cardiogenic shock and 0536 is those patients with STEMI or cardiogenic shock.

A few important distinctions with the previous measure that was just discussed to highlight. One of which is that the patients who die and are accounted for in the previous measure are not candidates for this

measure. This is death after discharge.

And the validation data that are presented are generated from matched claims data. Ultimately these will be matched with broader death records. So with the hopes of making them applicable to broader populations, so a modification of what I said before.

And I don't know if -- I think that's pretty much all I need to say. Again -- and they are intended for public reporting, right, and have been, you know, the models have been validated fairly extensively as you'll see in your materials.

DR. WINKLER: Fred, I just want to clarify. You said so this does not include the in-hospital deaths. Those patients are removed from this measure.

DR. MASSOUDI: Correct.

DR. WINKLER: So, this measure is only for patients who are discharged alive from the hospital and whatever else --

DR. MASSOUDI: Correct. Yes,

that's right. Intended to be complementary to each other and also to the previous measure.

DR. TING: Fred, I know my patients are sicker than yours.

So, this measure is I think going to be very similar. My secondary discussant is George Philippides. And I think the next measure is, as Fred pointed out, is going to be very similar. So, the methods are -- other than the patient population. So we'll probably have both go pretty quickly.

The description of this measure is 30-day all-cause risk-standardized mortality rate following PCI for patients without STEMI and without shock. So these are the lower-risk patients.

The level is at the hospital. As far as the evidence this is a health outcome risk-adjusted with NCDR CathPCI clinical registry data and linked to the CMS database for 30-day mortality.

There's certainly a processes of

care that are associated with improved outcomes. So, looking at the evidence algorithm it's actually a pass. Is that right? That means high evidence. It's an outcome measure.

Do we even vote on that? We do vote? Okay. But the algorithm says it would be a pass.

DR. GEORGE: Any discussion? All right, we'll vote on the importance.

MS. LUONG: The timer starts now. One is yes, two is no.

DR. WINKLER: This is a vote on the evidence for an outcome measure.

MS. LUONG: Can everyone just re-vote there? Thank you. Eighteen for yes, one for no.

DR. TING: The opportunity for improvement. The performance on this measure from 2010 and 2011 was 1 percent, 4.2 percent with a mean of 1.8 percent. So, it's 98.2 percent are surviving to 30 days but still

there is a gap of 1 percent to 4.2 percent. So I think that's a moderate opportunity for improvement.

There are no evidence for disparities based on proportion of African-American race or dual eligible patients.

MS. DELONG: I'm a little confused about these mortality rates. They seem consistent with the overall mortality rates for PCI. Are these exclusive of inpatient?

DR. MASSOUDI: So, you mean they're consistent with the in-hospital mortality rates?

MS. DELONG: Pretty much. And these are exclusive of --

DR. MASSOUDI: This is 30 days though after hospitalization. Thirty days after discharge.

DR. TING: These patients survived --

MS. DELONG: Thirty days after discharge.

DR. MASSOUDI: Right.

MS. DELONG: And they're as big or bigger than the inpatient. These rates you have here are excluding inpatient. They're after a live discharge.

DR. MASSOUDI: Correct. Yes, right.

MS. SLATTERY: So, just by way of reminder again. This is where we would like to emphasize the fact that these are always intended to be reported as a pair of measures. But particularly when talking about post discharge. These are always intended to be a pair of measures.

I don't know how when they got loaded into NQF's system they got numbered in the sequencing order. So you are looking at the parallel measure that's got the lower gap. But it was designed to avoid drift into the unknown. So patients suddenly not being identified as a STEMI or cardiogenic shock. So just by way of reminder.

DR. TING: Lara, was there any consideration just having one measure but just stratifying them as people who are -- just like we could report as one measure and then people who are low-risk versus high-risk as opposed to two separate measures?

MS. SLATTERY: Oh, yes. First go-around, sure, that was -- oh yes. And so what you have is two measures. That was determined as the best approach. Always intended to be reported as a pair.

DR. TING: Got it.

MS. DELONG: So the description of the measure says the numerator is -- the outcome for this measure is all-cause death within 30 days following a PCI procedure. It doesn't say following discharge.

DR. MASSOUDI: Well, it may be 30 days following the PCI procedure but it does not include patients who -- but it's patients who are discharged alive.

MS. DELONG: Yes, but --

MS. SLATTERY: So, I think you're correct that that's not as called out in the description. But in the specifications of the measure it does specifically state that patients must be discharged with status alive.

DR. VIDOVICH: So, just help me understand. So you're essentially doing landmark analysis, right? You're excluding the patients who died in the hospitalization. Then you restart the clock again.

Is there a specific reason to change the denominator to reduce?

MS. SLATTERY: So these pair of measures were developed after the in-hospital measure had been developed. The in-hospital measure was being reported systematically already. We are close to being able to start to implement these systematically.

There are some other considerations. It is not trivial to get at the post-procedure component. And so they were harmonized that way to pull it out and

allow them more specificity in the post procedure, knowing that the in-hospital procedure was already being reported out.

There are also again as you'll note in here some slight variables, particularly with STEMI and cardiogenic shock that were more significant in the post procedure than what we were seeing in the in-hospital.

MS. DELONG: So, I'm so confused. Is this 30 days post procedure or 30 days post discharge which could be 60 days post procedure if somebody were in the hospital 30 days.

MS. SLATTERY: It's 30 days post procedure assuming the patient was discharged with a status of alive.

DR. MASSOUDI: So it's parallel to what's used with the -- Joe, you can speak to this. But it's what's used -- is a similar process for what's used with STS for their 30-day post-bypass mortality. If I'm not

mistaken, Joe.

MS. DELONG: But those differences --

DR. MASSOUDI: It's discharge.

MS. DELONG: -- are very small. Post discharge to 30 days is usually very small from what I used to see in those databases.

DR. CLEVELAND: I'm not even sure we stratify post discharge. I think we just look at 30-day data.

DR. MASSOUDI: So, I'm not sure, Dr. DeLong, what the issue is. Can you please -- I mean, if the time from procedure to discharge is small I'm not sure what the --

MS. DELONG: When we looked at it I believe it was in the STS data set. And we looked at the difference between 30-day and in-hospital. It was minuscule. It was almost indistinguishable from in-hospital. So, what I'm saying is if you only look at that increment you may not get much signal.

DR. MASSOUDI: If you look at -- I'm sorry, which increment is that?

MS. DELONG: So, you've got a site that has a 2.1 percent in-hospital mortality. Their 30-day mortality might be 2.3. The increment that you're looking at, 0.02, or 0.2 is very, very small.

DR. MASSOUDI: But the increment is what you see in the data that are presented.

MS. DELONG: That's --

DR. MASSOUDI: The data that are presented exclude -- these are real data and they exclude the patients who died in hospital. That's the increment.

MS. DELONG: But they're almost the same numbers as we saw in the in-hospital mortality.

DR. MASSOUDI: That may be the case, but that is the incremental difference between the two. They happen to be similar but that is the incremental difference between

the two. They're not overlapping numbers.

DR. HOLLANDER: So, I want to follow up on that. Because in effectively every study ever published the event rates, the beginning are like this and go down. And so I know that's what the report says, but I question whether that's actually right.

Because I think Liz's take on it is probably right and consistent with every post-PCI study that's ever been done. Your events are early on and the further out you get the less likely events are.

It seems to me incredibly unusual to have near-similar event rates post discharge and in-hospital. And I just wonder if it's --

DR. MASSOUDI: Remember, though, that the time of ascertainment differs as well. So in-hospital tends to be a relatively short time frame. Right? And we're talking about 30 days. So it may be a declining rate, but it's over four or five times the period of

time of ascertainment.

And again, early is sort of in the eye of the beholder. In that yes, event rates may drop after the early period, but 30 days is relatively early in the context of an MI.

MS. SLATTERY: When they developed the measure initially for the last go-around also they actually looked at it all the way out to 45 days. You're right, most of the event occurred probably more around 21 days. But they made the decision that it was probably to go with the 30-day cut point than all the way out to 45 days.

DR. JAMES: I'm the sole vote that said no on this. And I know that right now that particularly with CMS the use of 30-day all-cause mortality or readmission rates or whatever is very popular.

But a number of us have had concerns about that it really should be a measure of something related to the procedure or the disease entity.

And I still have a problem with something. When you're getting to these small numbers, these incremental numbers of people who are discharged from the hospital, that's where the rate of being struck by an automobile, struck by lightning, or having something completely unrelated starts to go up as a percentage.

And so that's why I've got concerns with this one. And I'm sorry, also with yours.

DR. MASSOUDI: Yes, I think it's a reasonable point. I mean, there are a couple of issues to address there.

One of which is that it's a great idea to think about well, let's just look at procedurally related deaths. Put that in front of a committee of people and there's absolutely zero agreement on what constitutes procedurally related or not. I mean, short of an automobile accident. But even then, maybe someone had a syncope from a reinfarction.

And they drive their car into a tree and they die. So there's obviously, you know, it's sort of in the eye of the beholder in one.

And the other issue is that there are statistically distinguishable differences amongst sites. And so even though there's noise there there's reason to believe that there is variability, meaningful variability in mortality that is beyond the play of chance when you look at these at a site level.

DR. JAMES: There's also statistical arguments counter.

DR. MASSOUDI: Yes, of course. I mean we can -- yes.

MS. MITCHELL: I'm good. I was going to beat a dead horse. I just wanted clarification on the 30-day post procedure versus 30-day post discharge. And the clarification is that it's post discharge.

DR. GEORGE: Any other discussion before we vote on importance and opportunities?

MS. LUONG: The timer starts now. One for high, two for moderate, three for low and four for insufficient.

DR. WINKLER: You're voting on performance gap opportunity for improvement.

MS. LUONG: We have eight for high, six for moderate, five for low and two for insufficient.

DR. TING: So, priority is next. And so this cohort actually includes some sick patients, patients with non-acceleration myocardial infarction, patients with left main complex three-vessel disease, heart failure. The only people who are in 0536 that are actually STEMI and shock. So, it wouldn't -- I'm not surprised there are some deaths here.

As far as priority I mean I think we've talked about PCI multiple times already. This is a common procedure and mortality I think is an outcome that patients care about.

DR. GEORGE: Any discussion on importance? Priority.

MS. LUONG: For high priority the timer starts now. One is for high, two is for moderate, three is for low and four is for insufficient.

We have 15 for high, 3 for moderate, 2 for low and 1 for insufficient.

DR. TING: So for scientific acceptability we've talked about the numerator being -- this is an outcome measure for all-cause death within 30 days following the PCI procedure in patients without STEMI or shock.

At the time of the procedure the denominator includes all inpatients and outpatient hospital stays with a PCI procedure for patients at least 18 years of age.

Includes outpatient observational stay, patients who have undergone PCI but have chosen not to be admitted.

There are several denominator exclusions which all seem appropriate based on multiple procedures in the hospital, transfer patients, or low-volume sites. So, the -- and

the calculation of expected versus predicted mortality, observed versus predicted mortality is really based on 18 clinical variables within the NCDR database. So that seems solid in terms of scientific validity.

DR. GEORGE: Any discussion on that? If not we'll move onto reliability.

MS. LUONG: Reliability. Sorry.

DR. TING: So reliability testing was done both at the level of the performance measure score as well as the data elements.

It was the test/retest you've heard before. Each hospital had their data sets randomly selected into two data sets. The intra-class correlation coefficient was 0.256 which indicates fair or moderate agreement on reliability testing.

DR. GEORGE: Any discussion on reliability testing?

MS. DELONG: Did you look at the correlation between this and the previous measure? Is this not an appropriate time to

talk about this and the previous measure? Because it would be good to see how this one fares and actually -- it should be fairly consistent.

MS. SLATTERY: So, when they originally developed the measure, yes, they looked at correlation between the in-hospital and this measure being developed. I don't know that we revisited it for purposes of this measure project.

MS. DELONG: So, who developed this measure actually?

MS. SLATTERY: So, it was originally developed under contract with the Centers for Medicare and Medicaid Services. And then Yale Centers for Outcome Research and Evaluation was the analytic center. And then American College of Cardiology was also a partner on that.

When the measures originally went through the endorsement cycle CMS was listed as the measure steward. ACC is now taking

over measure stewardship with this project.

And so we allowed Yale access to all data to be able to take a look at and harmonize it with what was going on with the in-hospital measures as well.

MS. DELONG: Because I just recalled that we worked on a similar measure at DCRI.

MS. SLATTERY: At the time that measure actually was -- the in-hospital measure that you just discussed was originally developed by DCRI and as evidenced by it being the published papers coming out the lead authors are from DCRI.

But ACC is the steward and owner of those measures. So for purposes of development of this pair that information was provided to Yale.

DR. JAMES: I'm sorry to be so negative here. On page 38 the graphic there looks like a non-correlation. But am I looking at this thing wrong? I mean, I'm just

a country doctor.

(Laughter)

MS. ISIJOLA: We probably have different pagination because we don't have access to quite the same documentation that you have. So if you could give us a little more of a landmark.

DR. TING: I think, Tom, that is why it's fair to moderate. You know, the ICC of, what is it, 0.256 shows. It's not great, perfect, strong. Fair to moderate.

DR. MASSOUDI: I don't hunt squirrels so I don't know.

(Laughter)

DR. GEORGE: Any other discussion on reliability? If not we'll move to a vote on reliability.

MS. LUONG: The timer starts now for reliability. And it's one for high, two for moderate, three for low and four for insufficient.

Can everyone just point towards me

one more time? There we go. Thank you. So, 4 voted high, 11 voted moderate and 6 voted low.

DR. TING: So moving onto validity testing. This was done at the level of the data elements only. Overall agreement statistic was a median agreement 92 percent at the data element level.

I did have a question regarding was the validity testing done for all the data elements, or just the 18 that were in the model? That wasn't clear to me.

DR. MASSOUDI: Well, there's been broader validity testing of NCDR elements that goes through various cycles. But I think the validity testing that's addressed here is pertinent to the variables that are included in the model.

DR. TING: So there are 18 variables. And the median agreement was reported at 92 percent. And again, the measure is risk-adjusted using a hierarchical

logistic regression model with 16 risk factors.

And the calculated score is the ratio of predicted deaths to number of expected deaths multiplied by the national mortality rate. So it's very much like the AMI or CHF RMSR.

And let's see. The C statistic which is the area under the receiver operating curve was 0.807 for the validation sample which is considered acceptable. Anything above C statistic, anything above 0.7 is considered acceptable. So this was 0.8.

DR. GEORGE: Any discussion on validity?

MS. DELONG: So, Fred and I actually had -- Dr. Massoudi and I actually had a conversation awhile back.

And it would really be helpful I think, Fred actually suggested this, that there would be some templates for developers to use when they're reporting things like

this.

And one of the things that would be helpful. When they report percent agreement it's not necessarily meaningful. For example, on a data element -- I mean, you want to see data element-wise. But they could mostly be nos. So your agreement could be 80-90 percent. But where they disagree they almost disagree entirely. So you really want to see that 2 by 2 table of did they agree entirely. The problem is you need repeat measures for that.

DR. MASSOUDI: Again that's sort of a larger policy issue in terms of how the NQF wants to provide direction to measure developers.

But I concur that a certain degree of consistency around realistic standards that could be achieved and greater guidance, we're all for it.

DR. WINKLER: We're certainly open to the conversation to make things as

standardized and as easily understood for everyone as possible.

DR. GEORGE: Any other comments on the validity? If not we'll move to a vote.

MS. LUONG: The timer starts now. One is for high, two is for moderate, three is for low and four is for insufficient.

Ten voted high, nine for moderate and two for low.

DR. GEORGE: Feasibility.

DR. TING: So for feasibility the data source again is the NCDR CathPCI clinical registry which we've talked about for the other performance measures. For PCI the -- over 90 percent of hospitals that do PCI participate in this registry so I think it's quite feasible, consistent with the other measures that we've already looked at, endorsed.

DR. GEORGE: Any comments on feasibility?

DR. WINKLER: One comment that

always comes up and that's matching it to the CMS data. Is that a time lag in terms of being able to calculate the measure? What logistical issues do you encounter putting those data together?

MS. SLATTERY: So, just as a reminder, when we go to implement it will be based on CDC data, not CMS data.

But yes, there are some lag time issues, particularly in this instance because we are limited to going with CDC data.

That's the only avenue currently available to us. We are tracking regs to see if we will be able to get access to Social Security Administration vital status data which could be a significant game-changer in terms of timeliness of being able to report this out as well as frequency of being able to report this back to our sites.

When we initially put this measure forward and expressed the desire to be able to report it on all patients we did have access

to Social Security Administration master death file. During that time there have been some changes but there are new regulations being introduced to potentially create the opportunity that we could get access to that data.

So right now we're dealing with CDC data and there is a time lag there as well. We just sent off our data files for 2011 and 2012 so those will be able to be matched.

One of the other challenges. We had one time previously gone through the CDC process for applying. And it was for a different registry.

This is our first time sending off PCI data for the match. We are hoping that we don't encounter some of the same questions. Because we do find that different reviewers at CDC come back with different questions. So, there are challenges.

DR. GEORGE: Any other comments on

feasibility? If not we'll go to a vote.

MS. LUONG: The timer starts now for feasibility. One is for high, two is for moderate, three is for low and four is for insufficient.

Twelve voted for high, eight for moderate and one for low.

DR. GEORGE: Validity.

DR. TING: Use and usability. I don't want to go back to validity. So, usability and use.

This measure as I understand is currently not in use. But as far as historical trends or secular trends, from 2006 to 2008 the 30-day mortality rate was 1.4 percent as a median. Now it's from 2010 to 2011 that has increased to 1.8 percent. So, there would be some use to continuing following those trends and seeing if we can improve.

DR. GEORGE: Any comments on usability? If not we'll vote on usability.

MS. LUONG: The timer starts now. One is for high, two is for moderate, three is for low and four is for insufficient information.

Can everyone just do it one more time? Thank you. Nine for high, ten for moderate, one for low and one for insufficient information.

DR. GEORGE: All right. Any final comments before we vote on the overall?

DR. TING: Will we talk about competing measures? Or that's after this vote? Thank you.

DR. GEORGE: All right, we'll go to an up or down vote.

MS. LUONG: The timer starts now. One is for yes and two is for no.

Seventeen voted yes and four for no.

DR. TING: Terrific. The last comment I had was about competing measures. There are four other measures that look at 30-

day all-cause mortality. So that's for heart failure, acute myocardial infarction, COPD and pneumonia. So if any of those patients happen to get a PCI they would be in both sort of measures.

DR. KOTTKE: You get to vote twice. 0536, 30-day all-cause risk-standardized mortality rate following PCI for patients with STEMI or cardiogenic shock.

DR. MASSOUDI: Right, so this is sort of the teammate of the other measure. And I will say that probably the largest difference is the higher event rates in this population not surprisingly because it's STEMI and shock.

Just a small footnote. I have to catch a flight and will have to leave at 6:15. So I'm obviously not empowered to put anyone on the clock but I will have to leave at 6:15 which is fine. You have able representation here. But if I depart that's why. Thank you. Hopefully it won't be necessary.

DR. CLEVELAND: My nickname is Ted Cruz so I'll filibuster Fred for the next 23 hours.

(Laughter)

DR. CLEVELAND: Kidding. I thank Henry for taking a lot of the headway on this because this really is -- I'll be brief. Again, this is 30-day all-cause risk-standardized mortality following PCI. Really the difference between the previous measure is just these are sick patients, they truly are. So I think we can use the word "death" because it involves STEMI and cardiogenic shock.

Data source. Again, the NCDR, CathPCI. And again, the evidence or to skip ahead quickly to that. It's an outcome measure. There are data provided by the measure developer associating increased survival with the use of periprocedural clopidogrel, GP2B3 inhibitors. Participation continues quality improvement. So I found the evidence to basically meet the criteria to say

yes.

DR. KOTTKE: Kristi?

MS. MITCHELL: I have nothing to add.

MS. DELONG: Does the evidence speak to post discharge, or is it just 30 days post procedure?

DR. CLEVELAND: I guess the evidence is 30 days post procedure.

MS. DELONG: Within 30 days.

DR. CLEVELAND: Yes. Fred, do you want to amplify on that?

DR. MASSOUDI: Right. You mean in terms of the evidence-based therapies? Yes, correct, as Dr. Cleveland says.

DR. JAMES: In this case my objections are attenuated significantly because this is a group where the population at risk is much sicker, is more likely to have a cardiac event. So I'm going to reverse everything that I said on the prior one. I still believe in what I said before.

DR. KOTTKE: Any further discussion? So, vote on the evidence.

MS. LUONG: The timer starts now. One is yes, two is no.

Nineteen voted yes and one voted no.

DR. KOTTKE: Okay. Opportunity for improvement.

DR. CLEVELAND: So again, this -- as Tom pointed out, the spread here is quite high. The mean mortality in this is 12.6 percent, range 10.8 to 14.4. These are obviously 10 times what we saw in the two prior measures. So I think that there is a significant chance for improvement in those types of numbers.

DR. KOTTKE: Nothing? Any discussion? Okay, let's vote.

MS. LUONG: The timer starts now. One is for high, two is for moderate, three is for low and four is for insufficient.

Sixteen voted for high, four for

moderate and one for low.

DR. KOTTKE: Priority.

DR. CLEVELAND: Again, we can discuss what we said earlier. Coronary disease, PCI, STEMI and cardiogenic shock. I'd argue those are compelling priorities. Death.

(Laughter)

DR. KOTTKE: Well, anybody vote low priority or want to change their vote? Sorry. Okay, we'll just roll the vote over from the last one.

MS. LUONG: The timer starts now. One is high, two is moderate, three is low and four is insufficient.

Seventeen voted high and four for moderate.

DR. KOTTKE: Acceptability.

DR. CLEVELAND: So in regards to the acceptability the numerator statement again, all-cause death within 30 days following PCI. That's what's stated here in

the measure development. When patient with STEMI or cardiogenic shock at the time of the PCI.

There are some exclusions in the denominator. The denominators are exactly as what we saw in the last measure. The exclusions are PCI that follows a prior PCI in the same admission. That seems reasonable.

Patients with inconsistent or unknown vital status or other unreliable data. For example, someone who has a date of death preceding the PCI. Subsequent PCIs within 30 days to avoid double counting. And lastly, PCIs in patients with more than 10 days between the date of admission and the date of the PCI.

The argument was made that this is a rare, fairly heterogenous unusual situation, not well characterized. And I think I can accept that. So I think the exclusions are reasonable.

Again, the data source is the NCDR

linkage to this PCI registry with Medicare data.

I guess, we touched it on a little bit but I wouldn't underestimate the challenges in that linkage as we've already raised in discussion. So I really have no issues or concerns with the reliability.

I guess we can move onto reliability testing. So this was tested for reliability, both at the data element level and the measure score level.

Again, the reliability was actually -- the ICC for this was actually fairly low too, 0.122, slight agreement. So, when I follow it all the way, being a surgeon I can follow algorithms pretty well, I think what I arrived at was box 6B of the algorithm 2 which gives us a moderate reliability score.

DR. KOTTKE: Comments? Kristi, anything? No other motions? Vote.

MS. LUONG: The timer starts now. One is high, two is moderate, three is low and

four is insufficient.

Three voted high, seventeen for moderate and one for low.

DR. KOTTKE: Validity.

DR. CLEVELAND: In regards to validity, validity was tested at the data element level only. Again, data element validity so with the CathPCI as we've seen previously is fairly robust. Hospitals are audited, cases reviewed, the methodology is appropriate for that.

The only agreement statistic that was reported was a median agreement. Obviously there's no sensitivity or specificity around that.

However, in terms of potential threats this measure is risk-adjusted. It has a hierarchical logistic regression model with 13 risk factors. The C statistic for that was 0.83 with a validation sample which is quite acceptable.

There were a total of about 3,000

exclusions of the 40,000 patients or 48,000 patients looked at in calendar year 2010 to 2011 data set. Again, the exclusion criteria seemed appropriate.

So, I think that regarding validity -- oh, missing data. Less than 1 percent of the values are missing. And similar to the last discussion these values were imputed in a reasonable way. So validity I would rate as moderate as well.

DR. KOTTKE: Kristi, anything? Nothing. Any other? Seeing no action, let's vote.

MS. LUONG: The timer starts now. One is for high, two is for moderate, three is for low and four is for insufficient.

Seven voted high, thirteen for moderate and one for low.

DR. KOTTKE: Feasibility.

DR. CLEVELAND: Feasibility. Similar to the last discussion the data sources are registry elements with the CathPCI

and administrative data. I think we similarly discussed some of the challenges for that. I think it is feasible.

DR. KOTTKE: Kristi? Anybody else? Nobody? Let's vote.

MS. LUONG: The timer starts now. One for high, two for moderate, three for low and four for insufficient.

Can everyone just -- just keep pushing. Thank you. Ten for high and eleven for moderate.

DR. KOTTKE: Usability and use.

DR. CLEVELAND: In regards to usability this measure was originally NQF-endorsed in August of 2009. It is not currently in use. It is not publicly reported but there are plans for a phased implementation of public reporting. I think this is part of a rollout of kind of overall PCI mortality in the public reporting sphere.

One interesting note in terms of improvement. There's -- just as we saw with

the last data set when you looked at the 2006-2008 data set the mean risk-standardized mortality rate was 11 percent. That has increased to 12.6 percent in the 2010-2011 data set. So I think it does bear keeping an eye on the signal too as well.

Again, reasons for that a lot of things in terms of case mix, new addition to hospitals, et cetera, et cetera.

I suppose unintended consequences. This might be the one patient population that high-risk PCI patients would not receive PCI. It's always hard to know how people behave in a risk-averse type of thing with this. But I think the possibility does exist. I think we just need to be cognizant of that.

DR. KOTTKE: Kristi, anything? Nada? Any other comments? Let's vote on usability and use.

MS. LUONG: The timer starts now. One is for high, two is for moderate, three is for low and four is for insufficient

information.

Ten for high and eleven for moderate.

DR. KOTTKE: Let's vote on overall.

MS. LUONG: The timer starts now. One is for yes and two is for no.

For overall NQF endorsement 19 said yes and 2 said no.

DR. KOTTKE: So we're two minutes ahead of schedule.

(Laughter)

DR. WINKLER: The point you raised about intending to pair these measures. Actually, NQF can pair them in our system. And you could have chosen that when you submitted them but you didn't. But that's okay, we can retroactively do that. If the committee agrees that these are measures that should be paired.

And what pairing implies is that you do both of them. You don't do one or the

other or pick or do whatever you feel like. It's really the two together is a single entity. And in fact, we'll vote them that way. We'll put them out for vote. So they rise and fall together. They travel together. You report them together. And that's what pairing implies. And so it sounds like that's what the developers want. Does the committee agree that that's the way you would want to see these go forward?

I see --

DR. KOTTKE: Anybody disagree?

DR. WINKLER: Tom disagrees. Reason?

DR. JAMES: The reason is if the latter is a stronger measure and I think that it becomes weaker because of the statistical issues that I'm concerned about with the former one that would drag this one down.

DR. WINKLER: Is Tom the outlier? Does everybody else agree they should be paired? Anymore or Tom's our one outlier?

That's fine. That's fine. Okay, then we can put it forward that way.

DR. KOTTKE: We have a public comment period.

DR. WINKLER: Yes, we do want to do public comment.

MS. TIGHE: Cathy, if you could check and see if anyone on the line wants to provide a comment. And anyone in the room.

OPERATOR: Okay, if you would like to make a comment please press * then the number 1. There are no public comments from the phone line.

MS. TIGHE: Thank you.

DR. WINKLER: So, we will reconvene tomorrow morning to begin at 8 o'clock. We will have a continental breakfast available at 7:30. So we will be here and see you all tomorrow. Have a great evening. Thanks so much for your work today.

(Whereupon, the foregoing matter went off the record at 5:59 p.m.)

U

W

C E R T I F I C A T E

This is to certify that the foregoing transcript

In the matter of: Cardiovascular Measure Endoresment Project Standing Committee Meeting

Before: National Quality Forum

Date: 04-21-2014

Place: Washington, D.C.

was duly recorded and accurately transcribed under my direction; further, that said transcript is a true and accurate record of the proceedings.

Neal R Gross

Court Reporter