NATIONAL QUALITY FORUM
+ + + + +
CARDIOVASCULAR ENDORSEMENT
MAINTENANCE STEERING COMMITTEE
+ + + + +
WEDNESDAY, FEBRUARY 16, 2011
+ + + + +

The Cardiovascular Endorsement Maintenance Steering Committee met at the Conference Center of the American Immigration Lawyers Association, 1331 G Street, N.W., Washington, D.C., at 8:00 a.m., Mary George and Raymond Gibbons, Co-Chairs, presiding.

PRESENT:
MARY GEORGE, Co-Chair, MD, MSPH Centers for Disease Control and Prevention
RAYMOND GIBBONS, Co-Chair, MD Mayo Clinic
CAROL ALLRED, RN, National Coalition for Women of Heart Disease
ROCHELLE AYALA, MD, FACP, Memorial Healthcare System
SUNG HEE LESLIE CHO, MD, Cleveland Clinic
ANN DE VELASCO, RN, National Coalition for Women of Heart Disease
DIANNE JEWELL, PT, DPT, PhD, CCS, American Physical Therapy Association
DANA KING, MD, MPH, Medical University of South Carolina
BRUCE KOPLAN, MD, MPH, Brigham and Woman's Hospital
THOMAS KOTTKE, MD, MSPH, HeatlhPartners
DAVID MAGID, MD, MPH, Colorado Permanente Medical Group
GEORGE J. PHILIPPIDES, MD, FACC, Boston Medical Center
JON RASMUSSEN, PharmD, Kaiser Permanente - Colorado

DEVORAH RICH, PhD, Greater Detroit Area Health Council
ANDREA RUSSO, MD, Cooper University Hospital
MARK SANZ, MD, The International Heart Institute of Montana
SIDNEY C. SMITH, JR., MD, University of North Carolina at Chapel Hill
ROGER SNOW, MD, Commonwealth of Massachusetts
CHRISTINE STEARNS, MA, JD, New Jersey Business & Industry Association
KATHLEEN SZUMANSKI, RN, Emergency Nurses Association
SUMA THOMAS, MD, FACC, Lahey Clinic Medical Center
NQF STAFF:
HEIDI BOSSLEY, MSN, MBA
HELEN BURSTIN, MD, MPH
ANN HAMMERSMITH
ASHLEY MORSELL
KAREN PACE
KATHRYN STREETER
REVA WINKLER, MD, MPH
ALSO PRESENT:
SUSANNAH BERNHEIM, MD, Yale/YNNH Center for Outcomes Research & Evaluation*
JOHN BOTT, MSSW, MBA, AHRQ*
DALE BRATZLER, DO, MPH, Oklahoma Foundation for Medical Quality, Inc.*
SHERYL DAVIES, MA, Stanford University - AHRQ QI Development Team*
SUSAN FITZGERALD, RN, MBA, American College of Cardiology
JEFFREY GEPPERT, EdM, JD, Battelle Memorial Institute*
MARJORIE KING, MD, FACC, FACCVPR, American Association of Cardiovascular and Pulmonary Rehabilitation
HARLAN M. KRUMHOLZ, MD, Yale/YNNH Center for Outcomes Research & Evaluation*

ALSO PRESENT: (CONT.)

STEVEN LICHTMAN, EdD, FACCVPR, American Association of Cardiovascular and Pulmonary Rehabilitation*

KAREN R. LUI, RN, MS, FACCVPR, American Association of Cardiovascular and Pulmonary Rehabilitation*

KRISTYNE MCGUINN, MHS, American College of Cardiology

FREDERICK MASOUDI, MD, MSPH, American College of Cardiology

MATTHEW T. ROE, MD, MHS, Duke University Medical Center

ROBERT J. SCHMITZ, PhD, Mathematica Policy Research

RANDY THOMAS, MD, FACCVPR, American Association of Cardiovascular and Pulmonary Rehabilitation*

MARIAN V. WROBEL, PhD, Mathematica Policy Research

*Present via telephone

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

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

8:01 a.m.

CO-CHAIR GIBBONS: I presume everybody had an excellent evening pondering all the measures we'll be considering today. I don't see a lot of nods. Okay. So we're going to start off completing the measures dealing with AMI that we didn't get to yesterday, and I think much of our discussion around the measures towards the end of the afternoon has set the stage. So, hopefully, your memory will work short term overnight to recall most of the elements of that discussion as we proceed through these this morning.

So the first one is 287, median time to fibrinolysis, and I just want to point out that this came up in the discussion where Dr. Masoudi pointed out, I believe, that the elements involved in this one were very similar, if not identical, to 164, fibrinolysis therapy within 30 minutes. And as Kathleen goes through this maybe we can

take that into account. Kathleen?

MS. SZUMANSKI: I think we were beginning where we left off yesterday. Essentially, there is no significant difference between this measure and the other two that we discussed. And I know that David was interested in putting a proposal on the table that --

DR. MAGID: That we don't have to re-vote on each of the data elements because it's identical.

MS. SZUMANSKI: It is.

DR. MAGID: The data that goes into the measure is absolutely identical to what we've already voted on, and maybe we can just vote on the last one because really the only difference is they want to have an opportunity to present it as a median as opposed to a percentage. So it's just a different way of looking at the same data.

DR. WINKLER: Could I just hear from the committee your thoughts on do we need

both measures, is one preferable, is there an added value for having both measures if really you're taking the same data and just presenting it two different ways? Is there really a justification for having both?

DR. MAGID: Maybe we should ask the measure developer.

DR. BRATZLER: This is Dale Bratzler. Are you able to hear me? So, I mean, that's a conversation that came up in the initial development process, and I think the same issue applies to the inpatient setting.

CO-CHAIR GIBBONS: I'm sorry. You're breaking up a bit.

DR. BRATZLER: The median times are sometimes more actionable in terms of doing improvement. So I think that's the primary reason to have two separate measures, just reporting it two separate ways.

CO-CHAIR GIBBONS: We're having a little trouble. You're breaking up. If you

could think configuring your phone differently or whatever.

DR. BRATZLER: Yes, is this better?

CO-CHAIR GIBBONS: Yes.

DR. KOPLAN: I think Dr. Masoudi submitted this so maybe he --

DR. MASOUDI: Well, no, Dale can speak to it just as well. But I think what he's trying to say is that users have found the median time an easier target for quality improvement purposes and that the proportion facilitates comparisons among sites. And this is how it's been reported on Hospital Compare now for many years, and I think users find both of them useful based on our implementation.

CO-CHAIR GIBBONS: Sorry. Could you clarify that, Fred? Which is reported on Hospital Compare? Both?

DR. MASOUDI: Yes, both are reported on Hospital Compare, I believe.

DR. KOPLAN: You can't do both for one. You can't have one measure that says you collect both pieces of information.

DR. MASOUDI: No, it's the same exact measure, but what is recorded out for one -- so you have the whole population of people who got either PCI or fibrinolysis, and there are a group of times, and you can calculate those times either as a median of those times or the proportion of those times that fall under a certain threshold, and that's how it's done.

DR. RUSSO: Or I think maybe you mean like have one measure saying, like you could pick or, so you did this and you had it less than such and such or and make it one measure, so when you're reviewing it every time you're not reviewing two separate measures?

DR. KOPLAN: Yes. So like measure 287 collects both the median time and the proportion under 30 minutes. Like, say,

you're collecting your information but it's within one measure, or is that not something that gets done?

DR. WINKLER: I think that's something we can certainly recommend. Just hearing the discussion, it sounds like perhaps the median time is more useful to the audience who are actually on the ground looking at quality improvement and maybe less useful for the public reporting where the proportion might be more easily understood by broader audiences. And realizing that NQF's focus is really on the public reporting, you know, the question is, again, do we have more measures than we really need? Of course, your suggestion to combine them is another alternative to try and focus these and reduce the confusion.

CO-CHAIR GIBBONS: Can I ask the question and perhaps Dale on the phone can answer it? What would be the consequences of simply sunsetting the median time and just

reporting the percent?

DR. BRATZLER: So I think it has been pointed the data elements collected are identical. It's simply a matter of calculating the performance two different ways. So the hospitals actually do not collect two separate measures. They collect it once, and then we just calculate it two separate ways. I can't speak for CMS. I don't know if CMS is available on the call or anybody from the Hospital Quality Alliance that uses the data for public reporting. I think Fred and I have both pointed out that the median time was felt to be more useful for quality improvement, helping hospitals work on their targets for improvement. So I can't speak to CMS, though, what might happen if only one were, quote, endorsed.

CO-CHAIR GIBBONS: Well, for the moment, I think we have a proposal that we have the same voting pattern that we had yesterday and we'll have to come back to

reconcile the competitive nature of these two. Are there any objections to basically duplicating our votes from yesterday on the proportion measure for this median? Anybody who has an objection, please voice it at this time or --

DR. AYALA: Well, I think we may be able to decide on whether or not to drop the median one, especially because the healthcare system is going to make that calculation themselves and if they're only using that for operational purposes for process improvement then I don't know that it needs to be an indicator. So some of us might vote differently on the last question.

CO-CHAIR GIBBONS: Okay. I don't think they can make that calculation themselves, so we're clear. They would just then have the proportion as the target for quality improvement. Fred?

DR. MASOUDI: And the reason that it's not quite so simple for the institutions

to make the calculations themselves is that they can't necessarily know what the exact denominator is of their population, vis a vis the specifications, because that's coded off of billing data as opposed to all people who get fibrinolysis, for example. So you can't necessarily develop apples to apples comparisons off the same denominator. Again, I would just say emphatically this measure is, in terms of the institution, is absolutely no additional work, you know, in terms of the data collection. There's absolutely no difference in terms of what goes into this measure. It's just a difference in terms of how the output is reported in two different ways that have been found to be, in testing, useful in different ways.

CO-CHAIR GIBBONS: So I'm going to go back to the issue of can we duplicate the voting and let everybody ponder whether we want to return to this competitive issue as we deal with the whole issue of harmonization and

competitive issues because, otherwise, we're not going to get people on their planes today. So are there any objections to duplicating the vote?

(No response.)

CO-CHAIR GIBBONS: Okay. Hearing none, I think we will simply do that. And now, Kathleen, we've taken away most of your work. I'm sorry. We're going to now move the next measure, 290, median time to transfer.

MS. SZUMANSKI: This is another time-sensitive measure, and it relates to the time from recognition of MI to transfer to a facility that can provide acute coronary intervention. From the get-go, I think that we would recognize that this is probably important to measure, the fact that we would like to assure that patients get the level of care that is necessary, and if it cannot be provided in the facility that they first present at that they be transferred appropriately to the next site for appropriate

intervention. So the question is is this measure important enough to measure? I would recommend that it probably is.

CO-CHAIR GIBBONS: Questions? Discussion on this? Yes?

DR. BRATZLER: This is Dale. Can I make just a couple of just real quick points about this measure?

CO-CHAIR GIBBONS: Absolutely.

DR. BRATZLER: I think it's a very important measure. So it's important to remember that this is a really strictly defined denominator, but we think this has been one of the most important measures we've rolled out in the emergency department. So it is limited to patients with a diagnosis of acute myocardial infarction, ST segment elevation, or new left bundle branch block, just as the other acute MI measures are defined, only for patients who are transferred for cardiac intervention, so we specifically look at the record to see if they're being

transferred for acute cath lab, PCI angioplasty, words like that in the record.

So it's a limited population, and it's reported to the hospitals two different ways. We look at the overall median time to transfer so they can keep track of kind of how long it takes them to transfer all of their patients. But then we have a reporting population which is even more closely defined, and that is a patient who, again, has an acute MI transferred for acute coronary intervention and has no contraindications to fibrinolytic therapy. And the reason for that is, you know, I'm in Nebraska today so if there was a big snowstorm and a small facility has a patient that clearly has a contraindication to fibrinolytic therapy and so they make the decision to transfer the patient for PCI, we don't want to hold them accountable for the fact that it takes the ground ambulance a lot longer to get there than a helicopter in a snowstorm. So the reporting measure looks at

those patients for which the hospital had the choice of either giving fibrinolytic therapy or transferring the patient for PCI, and that is the reporting measure. It's a well-defined population.

DR. WINKLER: Just to clarify, Dale, this is Reva. I just want to clarify the specifications that you've presented to us are which of the two measures you've just described?

DR. BRATZLER: I'm sure it's the reporting measure, the one that's publically reported.

DR. WINKLER: Okay, thank you.

MS. SZUMANSKI: So the question would be is it important to measure.

DR. BRATZLER: I guess the only other point I'll make is that, you know, in a study from NRMI a few years ago the median time from arrival at the first hospital to intervention in the cath lab at the second hospital, the median time in NRMI was 180

minutes. So we really think that there needs to be a focus on making the transfers more efficiently.

CO-CHAIR GIBBONS: And, obviously, there are national projects devoted to that. All right. Let's take our first vote on importance of the measure, please. Are we software-equipped or not?

MS. PACE: No.

CO-CHAIR GIBBONS: We are not. So this is going to be the old-fashioned way. You'll actually have to raise your hand. Yes, all right. All who think yes for the importance of the measure. So just to make sure you counted correctly, all who vote no? I've been there, done that, and it's taken ten minutes to get the count right.

So let's move on to scientific acceptability.

MS. SZUMANSKI: I think that one of the challenges with this measure, and it is not particularly scientific, but the issue of

arrival time in today's emergency department is a challenge because we actually, in many places, because of the issue of crowding, have more than one time recorded. We may have the registration time, we may have the triage time, and we may have the medical screening exam time, and they are not necessarily the same. One would hope with this population of patients that they would fast track into the medical screening portion, but there is the issue of there may be a gap at the initial receiving end of the patients and the same gap may appear at the other end because of the simple issue of crowding that is touching every ED in the United States.

So that being said, I think that there is strong evidence to say that it is important to do this in a timely manner and that the patient should go to a site that can receive and provide the therapy, and there is evidence to show that this is a good thing to do for these patients who are presenting with

MI. As was mentioned yesterday, certainly, if you live in Montana and you have to go a distance in bad weather by some means, perhaps by air, perhaps by ground, that time is going to be, the clock is going to continue to tick even as you're arranging transportation.

So from a scientific standpoint, it does bear the challenge of evidence that we certainly feel that this therapy is urgent and necessary. Unfortunately, if you happen to have your MI in a location where the PCI is not available, you do have to be moved to another site of care. So that would be my summary kind of of the scientific background of this. There's kind of soft issues and hard issues that occur.

DR. BRATZLER: Sir, I'd like to respond just very briefly to that. And I think even if you happen to show up in a facility that maybe isn't quite as remote but there are issues around delays and the transfer or the other thing that we see

happening that we think is inappropriate is that the receiving center is telling the transferring center to withhold fibrinolytic therapy in a patient for whom there is no possibility of getting to the cath lab any where near 90 minutes. And so that's why for our reporting population we strictly define it as those patients who have no contraindication documented in the record to receiving fibrinolysis and who are specifically transferred for acute coronary intervention, a population for which the hospital, the transferring hospital has the choice of considering fibrinolytic therapy as the reperfusion strategy when they know there's no chance that they're going to get them to the next hospital within 90 minutes.

CO-CHAIR GIBBONS: Okay. Are there any other comments? And I'd just sort of highlight the point that Kathleen made about arrival time. We looked at the data yesterday and found that there was, roughly,

a 20-percent error rate in the arrival time when it was audited. Other comments about this before we vote?

DR. RUSSO: Just in terms of the arrival time, at least, and it may vary, but isn't that directly to the cath lab in most cases if they're going to have an acute intervention so there's no --

CO-CHAIR GIBBONS: No, this is arrival time at the initial hospital. So the outlying hospital, the transferring hospital, whatever you want to call them, this is the arrival time there. And then they're going to discharge, not discharge because they're never admitted, but they're going to then send the patient. So we looked at the arrival times and they're not very solid, as Kathleen --

DR. BRATZLER: They're not very consistent, but I think when you look at the variation in the numbers it's usually minutes.

CO-CHAIR GIBBONS: Unfortunately, I looked hard in the report, Dale, and I

didn't see that. Is that somewhere?

DR. BRATZLER: Yes, I think that's something we need to look at, but I'm pretty sure that's the case.

DR. SNOW: I'll just make the observation that, if you've got rotten times, one of the best ways to improve them is to have a measure that identifies them.

DR. SMITH: Are there demographic disparity matrices built into this? We know that, we're going to know that efforts will be made not only to transfer the mayor of a small town but all genders, ages, ethnic backgrounds promptly? Is that built into this?

MS. SZUMANSKI: It is not defined in the measure as such.

DR. BRATZLER: It is not defined. It can be. It is captured and can be calculated.

CO-CHAIR GIBBONS: All right. I think we want to go ahead and vote on scientific acceptability. Do we at least have

a slide that we can remind people what their choices are going to be in case they don't -- this is not yes/no now. You got to remember completely, partially, minimally, not at all or not applicable. Completely, partially, minimally, or not at all, all right? So show of hands --

DR. SMITH: Are we voting the disparities in?

CO-CHAIR GIBBONS: This is the total scientific acceptability, so that includes the disparities element.

DR. SMITH: They said it could be included. Will it be included?

DR. WINKLER: I think that's your recommendation that it should be included. It sounds like they can do it and you'll recommend --

DR. SMITH: So the vote would carry that? If I vote it's okay?

CO-CHAIR GIBBONS: Carry that recommendation to them. All right. So with

that recommendation to them, completely show of hands. Partially show of hands?

MS. SZUMANSKI: Eight.

CO-CHAIR GIBBONS: Just in case, minimally? Not at all? Moving on now to usability.

DR. WINKLER: I think this measure is currently in use. It is currently calculated and reported. Again, I don't see any additions that may make this measure clearer or not clearer other than the value of identifying discriminatory issues related to the population that is being served. It is harmonized. That's what I can say about it.

CO-CHAIR GIBBONS: All right, that's fine. Any other comments about usability? Okay. If not, we're going to move ahead and vote on this one. I'll remind you of your choices: completely, partially, minimally, not at all. So all those who say completely?

MS. SZUMANSKI: Thirteen.

CO-CHAIR GIBBONS: Partially?

MS. SZUMANSKI: Eight.

CO-CHAIR GIBBONS: Minimally? Not at all? I just realized we have to make sure no one is voting twice. And then, lastly, feasibility.

MS. SZUMANSKI: I think the feasibility issues, it can be collected. I think the challenge, again, is with the time issue and that clarification would certainly help the abstractors who are looking through charts trying to find out really what is the time they arrived and would certainly help this measure in the data collection standpoint. There are no e-specifications currently. Funding is pending, and it is hoped there will be some. It is easily collected. Abstractors will tell you, other than the data time, it's not too bad based upon some of the other measures that they collect.

CO-CHAIR GIBBONS: Other comments

on feasibility? If not, we're going to go ahead and vote. All those who believe it's been addressed completely? Partially?

DR. WINKLER: Twenty-one for partial.

CO-CHAIR GIBBONS: Minimally? Not at all?

DR. WINKLER: David, were you a minimal? Okay.

CO-CHAIR GIBBONS: Got to get those hands down faster. He was just making sure he got counted. All right. So now the final vote is does this measure meet the criteria for endorsement? A further discussion? This is yes/no. All that say yes? No? Okay. So we have now completed that discussion of 290. We're going to now move on to 0160, beta blocker prescribed at discharge. Devorah?

DR. RICH: Good morning. Okay. So the measure here is the beta blocker prescribed at discharge, and this would be the

percent of AMI patients prescribed beta blocker at hospital discharge, and this is a process measure. This is a really important measure in terms of reducing morbidity and mortality, and the ongoing use of this measure is designed to ensure high performance. When you look at the data on this, there is high performance. That is from looking second quarter `09 through first quarter `10, it ranges from 98.1 to 98.2. So there's high performance, and it does show disparities and even among the disparities, while there is definitely disparities, it's very close. It ranges from a low of 96.3 to a high of 98.3. So it's clustering pretty close together, but this is a strong predictor of overall ability to maintain health.

So are there questions? I mean, I think it's an important measure. You know, when I first looked at it, I was like someone said 100 percent, you know, why are we measuring this? But it doesn't show the, we

certainly don't know what the whole range of performance is. And in reading this over, it seems to indicate that by putting this forward as a measure it will make sure that it stays on the top of people's minds.

CO-CHAIR GIBBONS: Are there comments about importance? We discussed beta blockers quite a bit yesterday afternoon.

DR. KOTTKE: Ray, can I ask a quick question?

CO-CHAIR GIBBONS: Sure.

DR. KOTTKE: On indicators that have been decommissioned because they approach perfection, does anybody ever go back and look at the effect on social entropy on performance? In other words, you know, the old dust to dust, ashes to ashes thing that systems fall apart when you aren't watching them?

DR. WINKLER: I think, certainly, we've heard that voiced as a concern, but I don't actually, I'm not aware of any

particular, you know, rigorous review of evaluating that. Helen?

DR. BURSTIN: I've seen some early work from the VA suggesting that there wasn't an effect when measures were taken out of public reporting, but I've not seen anything else. But, again, it hasn't been published yet.

DR. RUSSO: And just for this measure, I think that this is such an important thing. Even if it were 99 percent instead of 98, I would think it might send the wrong message to not measure it at least right now. It's just so important.

DR. KING: If I can comment, I think that every one agrees it's important. That's why there's 98 or 99 percent compliance with it, just like with the aspirin of yesterday. And we've heard a plea and in our instructions to reviewers we were told to not maintain measures that were near perfection and yet we are continuing to do it. So if

we're trying to reduce the work of quality in the world and abstractors and poor little secretaries typing things into their Excel spreadsheets, it seems like the aspirin at 99 percent and this one at 98 and a half are two candidates, you know, to consider. And it's kind of like, when you cut yourself you get a tetanus shot. I don't know. That's not written anywhere and there's no national program, but everyone just still seems to remember that. I don't think there will be a significant problem because you're trying to use this as a quality improvement as something that discriminates, and this is a non-discriminating measure.

DR. RUSSO: Of maybe we could, you know, there's some question, I guess. It looks like there's more analysis in terms of disparities. Is there any disparity in terms of -- maybe not. And maybe eventually it should be retired. But I think if we do something like that, we should have some

consistent way of doing it and just come up with some standard to say if you're above 98 percent and there's no disparities. There should be some consistency in what we do if we start to retire measures. And it would be nice to see. I guess they're going to do some more analysis on the disparities, if there are; but I guess there may not be.

DR. RICH: You know, when we go to look at what we want to retire or not, I guess it would really depend on what's on a full plate. So I would recommend at this point keeping it. It's an important measure. I certainly agree with you, Dale, that, you know, it's almost at perfection, but if there's a huge difference between 98 and 100 percent. And if we believe we can get to 100 percent then maybe we want to keep it up there so people don't get the message that we don't care about it anymore.

MS. SZUMANSKI: Another way you might look at this is the percentage. If you

say two percent of 105,000 patients, how many patients is that? That's looking at it from a slightly different pair of glasses, but that's a significant number of patients in that cohort.

CO-CHAIR GIBBONS: The other part of that, though, is that when you're at 98 and a half percent, you're in the rounding area. Your confidence interval is probably going to put you so close to 100 percent that I don't know if you can claim those patients reliability. If we can't retire a measure when we're at 98 and a half percent, we really need to think about when we can. And if we're not going to do it here, then somebody should sit down and work on that as a problem because, otherwise, the measures just expand and expand, and I think that's a problem that NQF has recognized.

CO-CHAIR GEORGE: We do have other beta blocker measures in slightly different areas of care, but perhaps that whole issue

surrounding beta blocker measures needs to be harmonized into those that are most important.

DR. BRATZLER: This is Dale. I guess where I'm getting a bit confused here is the language that's being used about whether this committee is looking at, quote, retirement of measures or whether they're looking at re-endorsement of measures because I think the entire committee, you know, that I'm hearing feels that the measure is scientifically sound. It's a good measure. There's nothing scientifically wrong with the measure. The question is around implementation about whether it still makes sense to have hospitals capturing the data on the measure. Helen I certainly know is in the room, and Helen is well aware of the conversation extensive with CMS about when they decide to retire measures that are scientifically sound and is there any way to do that intermittent surveillance to make sure that we don't see backsliding.

So all those conversations are going on. I think the question for us has been does it make sense to at least keep some library of measures that are endorsed because they're scientifically sound from which, you know, can be used to pull from in the future, if needed. I know, Helen, you've talked about some new category of measure.

DR. RUSSO: Or maybe like the one from yesterday, using a composite. Would it make sense to consider the ACE inhibitor plus beta. They both function so well. Doing one of those composite measures, considering that. But this is one that everyone thinks of at discharge, all the guidelines, you know, AHA work. So this one is really high now because I think people have worked really hard, but maybe if we use it, because it is so important, combine it, like the smoking one yesterday with the lipids, so then people don't perform as well when you combine multiple drugs. Just something we might

consider.

DR. BURSTIN: Hi, Dale. It's Helen. It's something we've actually talked a lot about. We don't really have a category for this yet, although the evidence task force that just completed its work did specifically make a recommendation about measures that were fairly high level performance. You want to just review that, Karen?

MS. PACE: Basically, you know, as I said yesterday, they didn't feel they could set a particular threshold that is contextually driven, and certainly you experts around the table know more about this. But one of the key considerations is whether there are associated outcome measures that are being measured so that, you know, if this process is associated with one of the outcome measures we have then if you've got an outcome measure already, you know, very high levels of performance, do you need to continue endorsement of the associated process

measures. So I think that's something else to kind of throw into your discussion about that.

I don't know, you know, again, as Helen said, we haven't set up a separate category for measures that are sound but are, quote, topped out, but certainly, you know, if we don't continue endorsement, you know, the reason for that could be clearly indicated so that the measure could be used again. But it sounds like it would be used more from a monitoring perspective and quality improvement versus performance measurement, but something else for your consideration.

DR. KOTTKE: Ray, can I make a comment?

CO-CHAIR GIBBONS: Yes, Tom.

DR. KOTTKE: My quick spreadsheet calculations is that going from 98 to 100 percent would, in a population of 100,000 middle-aged adults, would prevent a tenth of a life lost per 100,000 people compared to if you use that same energy to make sure

everybody got the cardiac rehab it would be about 12 lives. And so it's really very marginal impact, so we have to decide how much opportunity cost are hospitals going to spend trying to get from 98 to 100 percent.

CO-CHAIR GIBBONS: Okay. Mark?

DR. SANZ: One last comment would be from the ground. When we've looked at our misses on this and Plavix after PCI, 100 percent were administrative error. So how much effort after a certain point should go into what sounds good as a clinical miss, missing giving someone a beta blocker, but in actuality is it didn't get reported in the discharge summary but when a nurse went back and called the patient they were on the beta blocker. It was an administrative miss. And when you get to these kind of numbers, most of the time that's what you're looking at.

CO-CHAIR GIBBONS: Helen?

DR. BURSTIN: I just checked, and this measure has already been retooled for

EHR, so that's another consideration is the question is how much of the actual searching needs to be done if this is clearly something put in in a discharge electronic system.

DR. RICH: Another thing to consider when we think about this for public reporting, you know, when you have a public report you can only put so many variables out there. And so what's of most benefit to the community to be looking at? You know, you want to really pick things that they're going to see, like, "Oh, here's where we stand, but we're stretching. We're trying to move forward," or, "We're doing great. If we all want to feel really good this is something definitely worth keeping in there."

CO-CHAIR GIBBONS: Okay. I think we've had great discussion, but we must move on. So we're going to take a series of votes on this one, and I would point out that if you're really concerned about this issue of is there enough room for improvement to justify

this effort, I believe, in terms of the way the process is set up, it should be reflected in the first vote on importance. So we're now going to take that vote on importance. So remind you, yes or no. So those who feel yes? Those who feel no? So the vote is 10 yes, 11 no. So we are done with this measure.

Next, 142, aspirin prescribed at discharge for AMI. George?

DR. PHILIPPIDES: We're in danger of having a very similar conversation. So just to jump right through the meat, this is post-MI with the codes of MI percentage of patients who received aspirin at hospital discharge. Exclusions are the ones that you would imagine: people who died before discharge or there were other important contraindications. So, obviously, this is a high-impact area. The point is going to be the one we just discussed: Over the last two years the numbers have really not budged, 98.3 to 98.5 percent, and it's hovering at about

98.5 percent. So it's not a particularly good discerner at this point, then, theoretically. It's obviously a very relevant outcome. I mean, clinically and scientifically, it's important. So I don't know if you want to have the same discussion and tweak the differences with aspirin or if you have any way in your spreadsheet --

CO-CHAIR GIBBONS: I think we're waiting for the Kottke model to do its magic. Dale, do you want to make any specific comments about aspirin at discharge?

DR. BRATZLER: No. This is actually Fred's measure. Again, my issue is, just like on the previous measure on beta blockers, if you de-endorse it and then five years from now we have issues around backsliding, you know, do we have to go through the entire re-endorsement process for a measure that's scientifically sound? To me, it's a question of implementation --

CO-CHAIR GIBBONS: I guess it's --

DR. BRATZLER: -- versus endorsement.

CO-CHAIR GIBBONS: I'll take a stab at answering that. I think the point has been made earlier that this is a matter of opportunity and how you apply the resources that you have in a system that we know has limited resources. Where you apply them are where you're going to get the most impact on actually benefitting patients. And applying a lot of resources to this one at this point in time I think is the question. Yes?

DR. JEWELL: So it looks like the disparities are the same or very similar in both measures, the one we just considered and the one that's on the table now. And I guess that piece still troubles me a little bit. I mean, if I'm in the Hispanic/Latino cohort, granted, 96 percent isn't a bad thing, but should we be worried about improving even two more percentage points for those people?

DR. PHILIPPIDES: That was one of

the two issues I was going to bring up before we voted, which is, is there evidence of disparity and if we lose this measure we lose the chance to look into that more closely. I think that, as a general rule, I've seen a greater need for that across many of these measures. The other thing is this one, as Helen points out, lends itself very well to electronic medical record review. It actually changed the way they look at exclusions. If you're given the aspirin, even though you have an exclusion documented, you're still sort of counting the credit, which makes it much easier to do an electronic search for the numbers. So it just makes the whole endeavor fairly easy. So I just want to bring up those two things, as well.

CO-CHAIR GIBBONS: Tom?

DR. KOTTKE: Yes, it's four times the impact of beta blockers at four-tenths of a life per 100,000 compared to about 14 for cardiac rehab.

DR. JEWELL: Starting at 98 percent.

CO-CHAIR GIBBONS: Four-tenths of a life per 100,000.

DR. JEWELL: So if you did 96 for the disparities. I mean, I guess the other thing about the disparities is that we're talking about populations who have other risk factors that are stronger in some cases in terms of their risk for co-morbidity and for complicated disease. So it's not only not getting the aspirin or not only not getting the beta blocker, it's the whole picture of how they present with their cardiovascular disease that, for me, just is troubling.

DR. KOTTKE: I think one thing to remember, I don't get what the guidelines or what the ACC registers are going to drop this just because we take it out as a, you know -- there's a lot of redundancy in the system besides just being in --

DR. SANZ: Well, speaking to

redundancy, there's six measures specifically using aspirin in their title, plus composite measures which I didn't go back to look at. So talk about redundancy. Aspirin is the most redundant measure of all of them, so do we really need, can we cut these down? It's not that aspirin is bad, but we have other measures that say if you have any coronary artery disease you should be on aspirin at discharge. Well, I submit that every MI patient has coronary artery disease and should be on aspirin at discharge and measured in the other group, so why do we need to do both? We have IVD, which I had not heard of until this group, but I submit that every one of these patients who's in that group. Why do we need to keep doing this as separate groups?

CO-CHAIR GIBBONS: Dr. Masoudi?

DR. MASOUDI: I have nothing to say.

CO-CHAIR GIBBONS: You're happy. Okay, good. I'm glad you're happy. So the

Chair is going to call -- Sid?

DR. SMITH: It's quick. Could we reassure that NQF will develop criteria that committees can use to determine whether something should be discontinued and also a plan for how they would be monitored and reinstated? It would be a real service to groups like this if, you know, even if you decide you want a sub task force from this committee, I'd be happy to work on that. But in the future, if we had objective criteria it would be very helpful.

DR. BURSTIN: It's a great point, and it's something that Evidence Task Force really grappled with and actually their decision was if there's an outcome measure and this measure is topped out do you need both, and the answer they said was no. So that was a question for you. Is there a clear outcome measure in the AMI realm, I think Tom is essentially doing that already by looking at mortality, that would suggest that if you have

the mortality measure and it's sound do you still need the process measure? And what the Evidence Task Force at least said was, no, that that's probably not necessary.

Now, the other thing to consider is that even if these are not endorsed as individual measures you will shortly have an opportunity to consider them in a composite. So there are ways to sort of keep it in ways beyond individual lots of measures being assessed regularly in public reporting.

CO-CHAIR GIBBONS: Bruce?

DR. KOPLAN: Just one last comment along what you're saying, whenever something gets no voted on, there's a clear expression that comes out the reason why, so it doesn't come out as if we as a group felt that it wasn't important. It was clearly expressed somehow that --

CO-CHAIR GIBBONS: There's not room for improvement. Okay. So the Chair is calling the question. We must move ahead if

we're ever going to get those rehab measures that have the bigger impact on lives that Tom keeps pointing out. So importance, yes, how many say yes? How many say no? Four yes, seventeen no. We're moving ahead to measure 137, ACE and ARB at discharge. Jon?

DR. RASMUSSEN: This measure is entitled ACE or ARB therapy for left ventricular systolic dysfunction for post AMI patients. It's measuring the percentage of AMI patients with left ventricular systolic dysfunction who are prescribed an ACE or an ARB by hospital discharge, so at any point in their hospitalization. For this measure, LVSD is defined as an ejection fraction of less than 40 or a narrative description of moderate or severe systolic dysfunction. It's a high-impact measure. The level of evidence is quite high with multiple randomized trials showing ACE inhibitors reduce morbidity and mortality in post MI patients with LVSD and ARBs have been shown to be a good alternative

for those patients who are not able to tolerate ACE inhibitors.

CO-CHAIR GIBBONS: Do you want to comment on the room for improvement issue, given our discussion thus far this morning?

DR. RASMUSSEN: Sure. So looking at the last few quarters, the national performance is around 96 percent. And using the numbers that they provide from 2300 hospitals, about 20,000 AMI patients, that gap represents about 800 patients.

DR. CHO: Can I ask a question? So the other criteria, like beta blockers and aspirin, it was like 100,000 patients. In here it's only 19,000 patients. Why is there such a disparity?

DR. RASMUSSEN: First, this is a subset of AMI patients, only those with LVSD. And I believe they have a public reporting program that looked at this measure, and it may be that the groups that reported out, that's where the --

DR. CHO: Because the other thing could be that they're not measuring LV systolic function.

DR. RASMUSSEN: Actually, that's an excellent point because when we get a little bit farther, looking at the exceptions, of those patients that were listed as having an exception, 64 percent of those were excluded because they couldn't find documented evidence of an EF or a narrative description of symptoms.

DR. CHO: So from looking at this data, I don't think this applies to the same argument as the aspirin or the beta blocker because of the lack of LV systolic function measurement for patients with MI.

DR. KOTTKE: About 5.2 lives.

CO-CHAIR GIBBONS: All right. So we have an estimate of 5.2. Dana?

DR. KING: Similar to our previous discussion, in my fog from yesterday I remember us discussing ACE and ARB therapy in

people with left ventricular function left ejection fraction of some 40 percent and number 0066. And that was for all people who had CAD that had either diabetes or left ventricular function injection fraction 40 percent. So now we have, but if you had a heart attack and less than 40 percent you need to get it, too. In other words, isn't this just a subset of 0066 yesterday which we endorsed and we said was fine and this one is, you know, and we need to consider whether it's necessary in addition to.

DR. RASMUSSEN: The anchor for this measure is a hospitalization. The measure 66 is on an outpatient basis, so they may have had an MI years ago. It's just that they were seen in that year. So this is a measure of inpatient performance. That's a measure of outpatient performance.

CO-CHAIR GIBBONS: And I would just point out that there is science. It's older science, but there is science that the

first three weeks of therapy makes a difference. Sid?

DR. SMITH: So I have a question that has just occurred to me, and that is assumptions that we make on samples. And an area of ignorance, that is I don't know what the denominator is for hospitals. But I look at the data here on patients, and there are, roughly, 2,300 hospitals surveyed, and we're told that they're at 96 percent. But when I look at the data for aspirin and beta blockers I see that 3,100 hospitals were surveyed. My concern is, to draw it to extremes to make a point, the 50 hospitals that are really interested in quality improvement show that 98 percent of their patients are getting aspirin, you could have 3,000 hospitals out there where only 10 percent are getting aspirin. And I'm concerned about drawing conclusions from a subset of hospitals that may not be truly representative if we really want to universally improve care. So the sample of

hospitals from which we draw data, it's very important to define that. And I'm just looking at these two sheets. I don't understand, I can understand why there would be fewer patients that would be candidates for ACE inhibitors than aspirin because not everybody has LV dysfunction. But I cannot understand why 2,300 patients report on ACE inhibitors and 3,200 report on aspirin.

CO-CHAIR GIBBONS: Fred, can you help us with that?

DR. MASOUDI: In terms of the relative differences in numbers of patients reporting for a different measure?

DR. SMITH: Not the differences. It's the number of hospitals reporting, the sampling. Your sampling 2,200 hospitals for ACE inhibitors and 3,200 for aspirin.

DR. BRATZLER: Yes. So this is Dale. There is no sampling at the hospitals. The data that you have I'm sure is all hospitals in the U.S. that voluntarily report

the metric. So how can I explain the disparity for ACE inhibitors and aspirin? Well, there are a whole lot more hospitals that take care of MIs that have very small sample sizes. Some of those hospitals may or may not have immediate availability of the ability to do left ventricular function testing. So I think I can pull out the data here in just one second and I'll cut in, but I can pretty well tell you that that is the total number of U.S. hospitals that are reporting the measure.

DR. SMITH: There's another bias that we got in to research in this area that those hospitals that do quality improvement have a bias to improving quality, and so their results, by definition, would be better than those who have no interest in improving quality or don't have the time or don't want to do it. So we need to know how many hospitals are out there that are not voluntarily involving themselves in this.

DR. BRATZLER: So I can tell you that remember that AMI has a much more limited number of hospitals that actually submit that data because they have low volume. They have either low volume or they don't admit the patients. I'm looking at second quarter 2010 data that comes into the QIO National Clinical Warehouse, which includes every reporting hospital that participates in the pay-for-reporting program which is about 98 percent of all eligible hospitals. There were 3,100 hospitals that submitted aspirin on arrival. There were 2,296 hospitals that submitted data on the ACE inhibitor measure. So this is just a difference in the number of hospitals that have eligible cases for the measures, but it is the total sample of all reporting U.S. hospitals.

CO-CHAIR GIBBONS: Sid, if it reassures you at all, to follow-up on the point that's been made, on the ground in southeastern Minnesota, I can make you a list

right now of six hospitals that would report the data on aspirin but would not report the data on ACE inhibitors who are within Mayo Health System. So with that as a sample, there's a lot of hospitals like them, and it's because they don't have LV function assessment and they have a very small numbers.

DR. SMITH: Well, the criteria for this are either an assessment of LV function objectively or that they had evidence for heart failure.

CO-CHAIR GIBBONS: Yes. And those six hospitals in southeastern Minnesota, the moment there's any suspicion, they're shipped to us, so that's what's happening here. You know, that's what happening. I'm saying that the sample, you're concerned that there's some, you know, hospitals that are doing this and not reporting, I don't think that's right. I think, as described, they had very small numbers, so they're falling off for that reason. If you go look at these hospitals on

Hospital Compare, there will be no data. They'll have that asterisk or whatever it is.

DR. SMITH: I'm concerned that a third of, roughly, 3,000 hospitals don't think it's important to report on a measure that significantly improves outcomes.

DR. MASOUDI: I don't think that's the issue. I think it's more the issue that they don't have the capacity to understand what the patients EF is and, therefore, they can't get --

DR. SMITH: I qualify that. And I'm also reassured that we apparently are getting 98 percent of the hospitals around the country involved. That's a good thing.

CO-CHAIR GIBBONS: Okay. I think we have to get moving, and we're going to vote on this. Importance of the question? We're back? We've got it? We're hoping. Okay. Now we have an opportunity to compare how long it takes to do it this way versus the old-fashioned way. I think we've demonstrated

that Ashley has a quick finger. All right, all right. ACE or ARBs, does it meet criteria for importance? Please vote. All right. Jon, any additional comments on scientific acceptability? Don't feel under pressure.

DR. RASMUSSEN: Just quickly. Numerator is number of patients prescribed an ACE or an ARB by discharge. Reliability, they select five cases per quarter across a number of measures from hospitals that have at least six discharges. Face validity regularly assessed by a technical expert panel. Exclusions are justified and are consistent with the other measures in the AMI set. I briefly mentioned looking at the exclusions. Almost 62 percent of those excluded were due to the fact that they did not have that documented EF or description of LV dysfunction. No risk adjustment reported. Meaningful differences quarterly benchmarked have been established. We've talked about that. Disparities, very small looking at race

disparities. High to low is only 1.4 percent.

CO-CHAIR GIBBONS: Other comments on this? All right. We're going to go ahead and vote on this. Okay. Eighteen completely, three partially. Usability.

DR. RASMUSSEN: Useful for public performance quality improvement is already a voluntarily reported initiative. Looking at other measures, this is the only inpatient ACE inhibitor. There are a couple of measures that are looking at ACEs and ARBs on an outpatient basis.

CO-CHAIR GIBBONS: I think we'll go ahead and vote on this one. Completely, partially, minimally, not at all. Nineteen to two. And, finally, feasibility.

DR. RASMUSSEN: The clinical data is generated during the care process. The exclusions do not require additional data sources. Susceptibility to error, they have a standard exclusion criteria added in the other reason for not taking ACEs and ARBs.

Trends do not suggest that that's a problem or is being gamed. Data collection, no evidence that it imposes an undue burden.

CO-CHAIR GIBBONS: Any other comments? All right. We'll vote on feasibility. We're unanimous on that. All right. And we're now going to move to does it qualify for endorsement. I'm sorry. That was a unanimous yes for those who are on the phone. Completely.

DR. RASMUSSEN: Just one comment that I have, and this goes across some of the other medication measures that we've looked at. With one exception, we're looking at 1.4 adherence for medications, and I certainly understand that it's a larger burden to track medication use over time, but I would encourage NQF in the future, as we start to see very high performance on these measures, the next step would be long-term adherence, 180 days outside of the hospital. As our data sources get better, I think that's a good way

to move in the future and a way to get away from we've got people who are 98 or 99 percent, stretch it out a little bit more shows better care, and then we can start to see some differences and places to improve.

CO-CHAIR GIBBONS: Point very well taken. All right. We're going to vote now on whether the measure meets criteria for endorsement. And that is a unanimous yes vote. All right. Thank you very much, Jon, for that expedited consideration of 137.

We're going to now move on to some true outcome measures. We mentioned earlier the importance of these, and we had a discussion yesterday about how you define certain criteria. I think there's little doubt that, from a public standpoint, things like mortality and readmission are the clearest outcome measures that any patient can understand and most of the public can understand and, from my experience, even people on Capitol Hill can understand. So,

therefore, we're now moving into that realm, and the first one is 230, AMI 30-day mortality. This is a re-up measure, and it's Tom Kottke. Tom?

DR. KOTTKE: Thanks. It's hospital 30-day all cause risk-stratified mortality rate.

CO-CHAIR GIBBONS: We're starting on 230.

DR. KOTTKE: 230 comes before 961.

CO-CHAIR GIBBONS: Okay. I'm sorry. There are too many multiple versions of the agenda, and I'm working off one of them, and we're going to do 230 first because it's the old measure. It would actually be wrong to do it the other way around.

DR. KOTTKE: Plus, Ray's the decider. Measure title is hospital 30-day all cause risk-stratified mortality rate, that's RSMR to the cognoscenti, following acute myocardial infarction hospitalization. The measure

estimates a hospital-level risk standardized mortality rate defined as death from any cause within 30 days after the index admission date for patients discharged from hospital. Mortality rates after MI are high. It's an important indicator. I don't think it needs more discussion than that for importance.

CO-CHAIR GIBBONS: Measure developers on the line?

DR. BERNHEIM: Yes. Hi, we're here.

CO-CHAIR GIBBONS: Good. Do you want to say anything further at this point before we begin voting on this measure?

DR. BERNHEIM: That was so nice and concise. We're happy to. Are there things that people want us to speak about a little bit more?

CO-CHAIR GIBBONS: We will see as we move through whether we have specific questions. We're going to start then voting

on importance. Don't vote yet. We're not ready.

DR. MASOUDI: Dr. Gibbons? There was a request from back here to have the individual who was speaking on the phone identify themselves.

CO-CHAIR GIBBONS: Thank you. Could the phone representative from CMS identify herself, please?

DR. BERNHEIM: Hi. This is Susannah Bernheim. I work at the Yale Center for Outcomes Research and Evaluation. We're the developer. And I have with me Elizabeth Drye and Kanchana Bhat as well from Yale CMS and Yale CORE.

CO-CHAIR GIBBONS: All right. Thank you very much. We're going to vote now on importance of the measure. All right. It's a unanimous yes, 19. We had two technical glitches. All right. Scientific acceptability. Tom?

DR. KOTTKE: The measure is

precise. It's 30-day all-cause mortality. The steward has demonstrated reliability in split-half analysis. There's a very nice analysis by the Yale group in circulation and some other stuff. Validity has been demonstrated by comparing the measure to a chart-based audit. The denominator exclusions are well defined and well documented. The frequency of the exclusions is documented in the accompanied 2010 Measures Maintenance Technical Report. The measure is fully risk adjusted with hierarchical general linear modeling. Risk-stratified mortality rate shows significant geographic variations that are clinically important. There are no comparable methods to measure for identification of disparities by gender and SES, but analysis indicates that these are small at the hospital level.

Now, there is one question about why not mortality rates strictly for coronary disease or for heart disease, and

the problem there is, one, you have to adjudicate cause of death; and, secondly, you can have changes in diagnosis, cause of death without actually changing death; and then, thirdly, it's pointed out more in the oncology literature the patients who receive chemotherapy within a week or a month of death, if everybody has an out here, they can basically order palliative care and then they're excluded from this measure. And so it will help prevent sort of the heroic interventions on people that are dying from other causes anyway, and so I think it's appropriate to have all-cause mortality rather than disease-specific mortality.

CO-CHAIR GIBBONS: Questions about scientific acceptability? I think, as Tom's pointed out, this has been extraordinarily well studied and vetted, and even if I didn't have people from Yale on the phone I would say that this is really a tribute to the group at Yale and the

leadership of Harlan Krumholz in doing that. Sid?

DR. SMITH: One question. Is this only involving Medicare patients, the database for this? Is there a mechanism by which younger patients are also involved? Are we talking about people over the age of 65?

CO-CHAIR GIBBONS: Correct. Over the age of 65.

DR. SMITH: So it seems like we would be well served by knowing what happened to people under the age of 65.

CO-CHAIR GIBBONS: Helen?

DR. KRUMHOLZ: Well, Sid, this is Harlan. I mean, I think that's a really good point. We're limited by the data source. We have every reason to believe that this measure would be just as true in younger patients, but we're searching for a data source that allows us to characterize the patients and look 30 days after

discharge. The national databases that exist in hospital you can do. Things with the managed care databases, but, of course, those are selected. So we're working very hard. I've talked to Helen about it. We're going to try to see what we can do to provide evidence that would allow a group to consider the expansion of the measure, but it's not for lack of interest or commitment to those groups but more about data availability at the current time.

DR. KOTTKE: If I could reinforce that, working for a managed care organization, we don't really know who's dead for at least 18 to 24 months afterwards because we have to wait for the state death, and I know Mayo is the same way. And it would be significant work to call people up and ask them if they're alive.

CO-CHAIR GIBBONS: Tom is exposing Minnesota's dirty laundry here. But he's right. It's somewhat

disillusioning to discover in our state that you really don't know for two years.

DR. KOTTKE: It is difficult, but that doesn't mean it's not important. I understand. I've been involved in similar work looking at heart failure patients. It's hard to get a handle on.

CO-CHAIR GIBBONS: By the way, welcome, Harlan. Okay. I think we're going to go ahead and vote on this, please. Scientific acceptability.

DR. RUSSO: Just as a quick question, is it Social Security Death Index or National Death Index? What is this? It's probably obvious to everyone, but how do you get the death quickly? Is it through Social Security Death Index or National Death Index? How do --

CO-CHAIR GIBBONS: Can somebody on the phone help us, please?

DR. BERNHEIM: Yes, sure. We actually get the information about mortality

from the Medicare Enrollment Database. As the measures are currently used, they're reporting on three years of data, but they're about a year delayed so we have plenty of time to have the full 30-day death information.

DR. RUSSO: So death then, it's not through Social Security Death Index?

DR. BERNHEIM: I believe that the Medicare Enrollment Database is updated through a different mechanism that also comes from the Social Security information. It's been shown to be an extremely valid and accurate way of identifying death in patients.

CO-CHAIR GIBBONS: Okay. So it was completely 19, partially 1. Moving on to usability. Tom?

DR. KOTTKE: The measure is already publically reported. The statistical adjustment method is the same one used for heart failure and pneumonia.

HRQ reports in-hospital mortality, but the 30-day mortality is independent of length of stay and cannot be influenced by cure decisions like early discharge if they look like they're not doing that well. And so I think it's completely.

CO-CHAIR GIBBONS: Comments or questions about usability? All right. Let's move on to vote. Eighteen completely, 2 partially. Feasibility. Tom?

DR. KOTTKE: The data are a byproduct of routine medical record coding. The data are available electronically. No additional data sources are required. The measure is already in use. It's prima facie evidence of feasibility.

CO-CHAIR GIBBONS: Comments or questions about that issue? All right. We will vote on feasibility. Somebody has a clicker that isn't working, but we have 20 completely. So now if I see 20 votes, I'll know somebody on this side of this table

voted twice. Okay. We'll now move on to the final vote. Does the measure meet the criteria for endorsement? So for those on the phone, that little side discussion reflected the fact that someone was not present, so our vote of 20 was, in fact, unanimous.

Unanimous vote of 18 to zero. I think I would like to reflect for the record that we've all recognized and several of you have commented privately to me that the submissions vary a lot in terms of how complete they are or how thorough they are and how easy they are to understand and read. I think this particular submission set an extraordinarily high standard. For those who didn't have the time to read it but want to learn about this methodology, as well as this measure, I would urge you to read the submission. I know that Dr. Kottke did an outstanding job of summarizing it that allowed us to work through it, but the

submission itself was truly of high quality.

We're now going to move on to the next measure, which is the hospital composite measure, 961, and Suma Thomas is the primary reviewer. Suma?

DR. THOMAS: So this is our first composite measure, and excuse me for my voice. So the primary objective is to summarize the measures for acute MI in a single composite that is useful, understandable, and acceptable to a variety of stakeholders. And there are other composite measures that AHRQ has produced in there in the appendix of other framework for composite measures.

It uses seven process measures and two outcomes measures for acute MI. The seven process measures we've mostly talked about: aspirin on arrival, aspirin at discharge, ACE or ARB with LV dysfunctions, smoking cessation counseling, beta blocker at discharge, fibrinolytics within 30

minutes, and PCI within 90 minutes. The outcome measures are the 30-day readmission and the 30-day mortality. The process measures are in a sub-composite, and the outcome measures are in a sub-composite.

The importance of a composite, there are three things that they pointed out: that the information from a number of composite measures can be summarized into a single measure, that the component measures can be aggregated at a level that's useful to consumers and providers, and that it can respond directly to patient-centered questions of which hospital should I go to. And the construct is a formative construct which is a combination of multiple measures intended to provide summary information. So each of the individual measures has been already found to be important, so this does seem to be an important thing to measure and report.

CO-CHAIR GIBBONS: Okay. Are

there other comments or questions about importance? Did the developers want to comment on this before we vote on importance?

DR. SCHMITZ: My name is Bob Schmitz. I work for Mathematica Policy Research. I thought that was a very nice summary, and I wanted to identify myself in case there were questions later. But, indeed, this is intended as a summary measure aimed at consumers. As someone noted just a few minutes ago, there are a multiplicity of outcome and process measures that consumers confront on Hospital Compare, and this is intended primarily as a means of summarizing that for consumers and is aimed primarily at them rather than at providers.

DR. RUSSO: And I think this just answers the question in terms of retirement of other measures. This is perfect for solving that problem.

CO-CHAIR GIBBONS: Sid?

DR. SMITH: I like the idea behind this measure of beginning to think in terms of how the system fits together. I wonder if we're talking to consumers about how you choose your hospital, and you have an MI you don't have a lot of choice in many instances. You go where they take you. And if we're looking at a 30-day mortality, not hospital, there could be elements that relate to what happens after the patient leaves the hospital that is not the result of quality in that hospital. And so if the intent is to advertise to consumers about a hospital, it seems that this measure is pulling in more than what's happening at the hospital. And I will say that I firmly believe and am passionately involved in the fact that hospitals' work do not end when the patients leave the hospital, and a marker for a good hospital system is having relationship with referring physicians and so forth. And, I mean, that's what we do,

and compliance with medical therapy at six months and so forth is important. But here somebody drops dead, gets hauled in, resuscitated, gets out of the hospital, I don't know that they're going to be going back to a setting that necessarily reflects that hospital system. And if the intent is to advertise this to the public, I just want to be sure that the hospital that's being advertised actually is in a way to influence all the parameters upon which it would be judged.

DR. SCHMITZ: I guess I would say that that 30-day measure is currently reported on Hospital Compare as an endorsed measure of, as one of the endorsed measures of quality. So it's been pulled into this composite on that basis.

DR. SMITH: There's probably not that many patients that would, I mean I think most of them live in the area, but I'm just sort of raising a question about that.

Again, following my strong statement in favor that this is the type of composite measure we need to be looking at.

CO-CHAIR GIBBONS: Others might want to comment, and Tom might want to comment from a data standpoint. But at least from a practice standpoint, I would argue that a lot of what goes on over the next 30 days is, in fact, the hospital's responsibility.

DR. SMITH: Well, I said that's for sure.

CO-CHAIR GIBBONS: I know, but I think it's reasonable for the public to see a measure that holds them accountable for that first 30 days because they can do a better job throughout our healthcare system of communicating the handoff to the physicians, and that certainly surfaced loud and clear from the heart failure mortality data when it first became apparent where, for the most part, hospitals did not realize

how many patients were dying between discharge and 30 days. So I think we have to do a better job as a healthcare system of handoffs and coordination, and there are various reasons, including payment system, why that isn't done. But, Sid, I think this is going to accomplish what it's trying to accomplish.

DR. BURSTIN: Let me just make one broad comment from the NQF perspective. We have definitely seen a push and a move towards measures of shared accountability. No one expects readmissions are solely on the back of hospitals or the receiving clinicians, but until we have measures like that it's going to be hard to really have that happen effectively. So this has been a measure that's gone forward. I think there's some really strong evidence that by having it be all cause and 30 day and requiring that interaction, we've actually seen some improvement.

DR. SMITH: Well, that's reassuring. Again, as I said, I strongly endorse this, passionately am involved in my own practice in what happens outside, but I live in the luxury of a system where I have electronic records that reach out miles, 30 or 40 miles to referring physicians. And I think Ray is in similar position. It's a little bit easier for us to really get in and affect that, but maybe that's an argument that other people ought to have that same luxury. And it sounds like you're making progress, so, again, I'm speaking in favor of this.

CO-CHAIR GIBBONS: Just for the public record, I really would like to point out that even systems with good electronic medical records, they do have patients who live in very rural locations with no electronic records, and that certainly includes us in Rochester, Minnesota.

DR. SMITH: We're linked in out

there, but, again, I would agree it takes more than a computer to take care of a patient. So it's the physician involvement that's key.

DR. MAGID: I'd add that this is consistent with the whole concept of the accountable care organization.

CO-CHAIR GIBBONS: That's an excellent point. Thank you, David.

DR. SANZ: I have a question regarding how or who would be reporting this since the purpose is public reporting. I mean, the measures, composites are excellent. Nobody is arguing. But is this going to go on some proprietary website? I don't know who Mathematica is either. I guess that's my question. I mean, am I going to be looking at a Google set of ads or my hospital and your hospital?

DR. SCHMITZ: We're CMS' contractor. This would appear on Hospital Compare.

DR. SANZ: Okay, okay. Thank you. And by the way, we are fourth lowest heart failure readmission in the United States at our hospital, and we don't have electronic medical records. So it isn't all there is. We do have to call the referring doc immediately after discharge, so there are other ways to communicate.

DR. JEWELL: I have a process question for the NQF staff. On a previous panel, I remember you all expressing a bit of concern and the panel was moving not to endorse individual measures that were then later rolled up into composite measures. I didn't hear you express that concern this last go-around, so am I safe in assuming that that has evolved?

DR. BURSTIN: I think we'll have to get to, I mean I think there are some differences there. So for example, NQF has a policy that all measures within a composite should be fully evaluated by a

committee. They don't have to be endorsed as stand-alone measures, which I think reflects back potentially on our last discussion of a couple of those process measures that are in here didn't reach the level of importance you thought as stand-alone measures. You would need to decide do they add value and should be part of a composite. We will get into some specific issues as you get into the elements within the composite. For example, there's one measure within the composite, smoking cessation, that was removed from endorsement not just because it was topped out, because it was not thought to be a valid indicator of smoking cessation in hospitals. So I think that's a slightly different issue than saying is it okay to have in a composite. I think you'll need to work those issues through as you get into the meat of it, but I think that gets beyond importance.

CO-CHAIR GIBBONS: All right.

Let's vote on importance to measure, please. All right. That's a unanimous yes. Scientific acceptability. Suma?

DR. THOMAS: So for the scientific acceptability, as I mentioned, there are the seven hospital process of care indicators and the two outcome of care indicators, which are sub-composites. In terms of the numerator, the numerator is a sum of all successes for acute MI process of care indicators which is weighted by one-half the reciprocal of the share of opportunities represented by acute MI process of care indicators and total opportunities, plus the sum of all successes for acute MI outcome of care indicators weighted by one-half the reciprocal of the share of opportunities represented by acute MI outcome of care indicators and total opportunities. The denominator is the total number of opportunities for success on all acute MI indicators used in the composite.

One comment I want to make is that right away I found this complicated, and the other composite measures that have been endorsed by AHRQ that I looked at in the appendix were, in my opinion, a little bit easier to understand. So I think the group's opinion about the scientific acceptability and the complexity of this will be very important.

Some of the decisions made in the methods were that they use values, not ranks, to decrease the likelihood of small differences in performance leading to large differences in the rank composite score. They imputed values for missing indicators, so composites were defining as many hospitals as possible. They adjusted the individual measures for reliability so that they avoided extreme variations for small hospitals due to random variation, and they used denominator weighting so the composite places more weight on measures that are

reported for relatively more patients nationally.

So the type of scoring that they used was called absolute scoring index with reliability weighting. And I think Mathematica can go into details if we need that probably.

And then in terms of missing composite scores, so you had to have at least four out of seven process of care indicators or one out of two outcome of care indicators or you were excluded. In terms of missing scores, if you met that criteria they would use the national mean to estimate the missing process of care or outcome of care value that was missing.

One of the things that I found important was that when they did testing of the missing scores, they found that four plus were missing in 35.7 percent of the time. Both outcome of care were missing 42.1 percent of the time, and all were

missing in 23.9 percent of the time.

In terms of reliability, it was based on the reliability of the component scores and they did do validity testing, as well. Another important thing that I found was that CMS has not decided how they will use it. When they talked about discrimination performance, they spoke about the Hospital Compare site and using better than hospitals, no difference than hospitals, and worse than hospitals, but clarify that they were not necessarily going to be using the measure in that way, which I found a little bit concerning not knowing how this measure was going to be used. And I'm not sure if it fits in here, but my general concern about this is that it's great to take things and make them one and make them easy for consumers, but you have to also deal with what that's going to mean to the provider. I know that this is important for consumers, and that's what

some of our goal is. But I also think about the providers and how they're going to look at this, and if you did take it and use it in such a basic way what kind of impact is that going to have down the line. We're all judged now more and more and more. It's important, but if it's this complicated can an individual provider out there, Joe Cardiologist, understand this? That's one of my major concerns because you're supposed to be able to break it down yourself and understand it. I'm not sure about that. That's one of my major concerns.

CO-CHAIR GIBBONS: Okay. I think we'll refer that question to the developers. Will providers understand this?

DR. SCHMITZ: That's the question? Will providers understand it? Well, I think we can put this together in a way that is, in fact, very easy to understand. The issue, there are two elements of the presentation in the form

that have a mathematical structure that makes it look quite forbidding, and much of this is done with the reliability weighting component. The reliability weighting is applied to the process measures to put them on the same footing as the outcome measures that were drawn directly from the Yale measures that you considered here. So they all are weighted in such a way that the mean is pushed toward the national mean for smaller hospitals, more so for smaller hospitals and less so for larger hospitals.

That said, once that calculation is done, the steps in the process can be made, I feel, quite easy to understand. And, in fact, the use of the national mean rather than other perhaps somewhat more accurate means of imputation was done precisely to make it easier for hospitals to take their own data and construct their value from it so that, in the end, part of the process here is for us to give each

hospital a set of instructions that allows them to take their data from Hospital Compare and create their own composite score from it.

CO-CHAIR GIBBONS: David?

DR. MAGID: So, I mean, it sounds like this is very well constructed and, from a statistical standpoint, very sound, though I'm not sure all of us with our sort of basic statistic background can appreciate that. So I wonder whether our colleagues at NQF can comment and at least say, yes, this is statistically sound. That would reassure me and we wouldn't have to spend as much time on that.

MS. PACE: This is Karen, and Helen and Reva had asked me to review the submission, so I have some observations and questions, actually, to pose to you and the measure developer. First of all, you know, I think it was very thoroughly described in terms of the methods and the analyses that

were done, so we appreciate the amount of detail and information that was provided. And I think one of the things that I noticed and certainly, you know, the way our criteria are worded would kind of lead us this way about some of the analyses that were done were more from the psychometric analysis, the intercorrelations, and the internal consistency which when you look at those you might question whether it is a sound measure. But I think and I'll pose this to the developers, I think we shouldn't necessarily rely on those to make that conclusion because I think this was constructed from the standpoint of using what's already there and trying to do a summary of it rather than starting with some conceptual model of what quality of care for AMI patients necessarily is and then creating a scale relative to that. And I don't know, Bob, if you want to comment on that.

DR. SCHMITZ: If it were appropriate, I would be jumping up. Yes, exactly. The psychometric analyses that are reported there are actually appropriate for a reflective measure that attempts to extract information from the individual measures and come up with a larger measure, similar, say, to the measure of IQ. We have explicitly abjured that in favor of the formative approach, which is make it an easy to understand, with apologies, summary that is explicitly intended as a formative summary of the measures. That's correct.

MS. PACE: Okay. So I think then the question for the committee is does this reflect an accurate summary of those measures that are already there and something that would be usable? So is the content basically sound, from your expertise in the area? I think one of the things that was pointed out and that I would have a question about is the seemingly large amount

of missing data, and I know this was touched on with earlier discussion. And so I'm not sure what the reason for that is for so much missing data. Is that a function of this being voluntary reporting?

DR. SCHMITZ: It is a function of reporting of hospitals to construct the measures that appear on Hospital Compare. So the reasons, the explicit reasons why hospitals might not report data for some process measures are a matter of conjecture to us.

MS. PACE: Right. But that's what I mean. It's not really, it's not a function of they've reported but the cases were too small to appear on Hospital Compare. It's that they haven't reported the data at all for the process measures?

DR. WROBEL: I'm sorry if Bob was about to answer this, but I'm Marian Wrobel, and I worked with Bob on the composite. For the mortality and readmission measures,

Hospital Compare does impose a bar of N equals 25 and, therefore, given that this is a summary of measures on Hospital Compare, we treat those indicators as missing. So one source of missing data is what you suggested: small hospitals being knocked out.

MS. PACE: That would be for the outcome measures? Because CMS has all the data they need to compute those outcome measures.

DR. WROBEL: Right, right.

MS. PACE: But you're saying that even though CMS has a rate for a small hospital that could be reliability adjusted, as you've talked about, you've just gone ahead and put in the mean for that, rather than starting with that base rate for those --

DR. WROBEL: Yes. And let me say two things about the reason for that. The first is that if we used rates that weren't

publically reported in the composite, a sophisticated user could back those rates back out by taking the composite apart. So if CMS' goal is not to publically report those indicators because they're not reliable, then it's necessary not to use them in the composite. A second thing about missing data is, of course, the other option would be to compute, to raise the standard for how much data must be available in order to compute the composite. And CMS, early in the process stated that an objective for this measure was to have it defined for as many hospitals as possible and we, therefore, needed a method that would define it for the majority of hospitals while still giving accurate signals about what is truly known about performance.

DR. THOMAS: Just one quick point. I said a lot, so I'm not sure if you guys caught this. To me, when you're missing, in 23.9 percent we're missing all

of the scores, it just seems like maybe this may not be our best, you know, composite if 23.9 percent were missing all of the scores, if you guys have some comment to that.

DR. SCHMITZ: For hospitals that are missing all the scores, no composite is created. And for those hospitals also no other measure appears on Hospital Compare.

DR. THOMAS: Right, I understand that. But to me it seems like it may not be the best composite if we have 24 percent of hospitals essentially not being able to use this to compare to other hospitals. I don't know. Maybe I'm looking at it in a different way.

DR. SCHMITZ: Well, there's a tradeoff here between imposing a requirement, which we did impose on ourselves that this was to be a composite of endorsed measures that appear on Hospital Compare, and having perhaps some larger number of hospitals represented with some

other measure. We didn't really know what those other measures would be. So our strategy really or our decision was to live with the fact that a substantial fraction of hospitals will be missing, given that that's a result for the endorsed measures in general.

DR. KOTTKE: I don't see that as a problem because small hospitals, those kind of things. If they're not graded, they're not graded. It's like you can't grade me on my angioplasties because I don't do them. And for those hospitals that simply don't have a large number of cases throughout, and I think we need to, the composite measure has to include the important variables that when you treat MI, those are important.

CO-CHAIR GIBBONS: So I would point out, you know, if you look at the key paragraph, that the outcome variables, the mortality and readmission variables missing

on 40 percent of the hospitals. So if I understood correctly, they're going to be imputed at the national mean; is that correct?

DR. SCHMITZ: That's correct if one of them is present. If they're both missing, there's no composite to compare --

CO-CHAIR GIBBONS: Right. So if one of them is present, the other one is going to be computed, and that's going to be a substantial percentage of the hospitals in this composite.

DR. SCHMITZ: Right.

CO-CHAIR GIBBONS: Okay. So we could have this discussion forever. I think we want to now vote on the importance of this measure as submitted. Scientific acceptability of this measure as submitted, now that we've had this discussion on imputation and missing variables.

DR. JEWELL: If I can, there's also the issue that I think both Suma and

Dr. Smith raised earlier about who is it being used for and what's it reflecting. So I appreciate what you said earlier, Karen, about the statistical psychometric approach to developing versus a more theoretical approach, and that's what you've done. You've pulled the endorsed measures, but not all of the elements perform that well. So I'm just curious if you re-examine the model more than once with pulling some of the individual measures out, like say the smoking cessation one which apparently didn't hold up. It's the same one, right, that you were alluding to? So we didn't, apparently NQF didn't or the committee at the time didn't find that one to be valid. So I'm a little concerned about what's in there, even though, theoretically, I get that they're all endorsed -- well, they're not all endorsed measures.

DR. SCHMITZ: I would say had we adopted the reflective strategy here, that

is a fully psychometric approach, and the approach of trying to develop a measure that was a truly reflective measure with optimal psychometric properties, I think that's the way, that's the kind of strategy we would have followed. But given that we were proceeding with the goal of developing a formative measure and that it was explicitly to summarize the measures that appear on Hospital Compare, that led us to switch away from that approach.

MS. PACE: Right. But one of our criteria is that the components either be, whatever approach you're using, either be NQF endorsed or evaluated to meet NQF criteria. So I think it's still a question regarding that.

DR. WROBEL: I do want to say another thing about the design of the measure, which is this is really intended to be a flexible methodology so that the composite will evolve as new measures are

brought on to Hospital Compare and as other measures are retired. So although we have written it up around the measures that are there now, the intended use --

MS. PACE: And another point on that is NQF can only endorse a measure that's specified. So even though this methodology is part and parcel of the measure, we aren't endorsing the methodology separate from the specific measure. So the other thing I just wanted to point out and make sure that I understood that your table 2K.3.1 gives the distribution of composite scores. So the 25th percentile score is 83.5 and the 75th percentile score is 84.98; is that correct? So that's another thing just for the discussion of the committee. This is a score that could be on a scale from zero to 100, and that's the distribution.

DR. AYALA: But isn't that a factor of the fact that they're plugging in

the mean for the missing components? That's the part that I don't think we've heard enough to help us understand how accurately that reflects the real performance of the hospitals.

MS. PACE: Well, that's the question. Right.

DR. SCHMITZ: It does in some regard result from that use of the mean. The use of the mean, to reiterate, was, in part, to make it easy to understand for hospitals and for them to recreate their own and also for us to be consistent with composite measures that have been developed by AHRQ and that were already endorsed by NQF so that part of the strategy here was to maintain some consistency across measures, composites that have already been endorsed.

DR. WROBEL: And CMS sought to distinguish between presenting the measure and presenting a strategy for display, although I do understand Dr. Thomas' point

about it's a lot easier to think about a measure when you understand how it will be displayed. This type of method, which is the AHRQ method, and this is how the 30-day mortality and readmission measures are treated, too, typically the display is hospitals are grouped into no different than the national mean, better than the national mean, or worse than the national mean so that the imputation process and the reliability adjustment for small hospitals is pulling them into that no different than the mean group.

CO-CHAIR GIBBONS: We have nicely had pointed out to us the importance of table 2K.3.1. It's in Appendix A, which I, for one, had a whole lot of difficulty finding, but my co-chair has found it. And if we have a way to show that, that would be wonderful because I think it will influence substantially the committee's view of this measure. It's on Appendix A. What page on

Appendix A?

MS. PACE: Page 20.

CO-CHAIR GIBBONS: Twenty of Appendix A.

MS. PACE: One other question for the developers regarding missing information. I know you chose to do denominator weighting, which means that the measures that are reported most often get more weight. So it's not necessarily reflective of items that are most important to patients or that all -- I mean, if most of these things said that all patients with MIs should receive, pretty similar, is weighting it by the amount of reporting really the best way to go, or what was your decision regarding that?

DR. SCHMITZ: Well, there are really two choices to make. One was to weight them all equally and another was to weight by denominator, the way we selected. Perhaps, in some ideal world, there would be

a means of weighting according to some measure of clinical importance or patient importance, but we wouldn't do that. Using denominator weighting has the effect of not necessarily minimizing the variance, but it reduces the variance relative to equal weighting. So it does tend to create a measure that is somewhat more precise than equal weighting. It is also an approach that's been used by AHRQ, so, again, we were using the principle of consistency.

MS. PACE: I think one of the distinctions between the AHRQ measures is that their components are different conditions, so they may have a mortality measure for procedures and those are each different procedures or mortality measure for conditions and those are each different conditions. But this is all for AMI, so it was just a question of decision.

CO-CHAIR GIBBONS: So if we could just, I think it will be hard for those at

the far end of the room to see these numbers, and I have to get my bifocals out to actually read these numbers. But the 1st percentile is 79, the 10th percentile is 81, the 90th percentile is 84, and the 99th percentile is 85. Thus, the spread from the 1st percentile to the 99th percentile is six percentage points.

MS. PACE: And just to point out, CMS is proposing the differential weighting so --

CO-CHAIR GIBBONS: Oh, so we'll go to the right hand column, which doesn't change things very much. First, the 99th is 79 to 86. That's a total of seven percentage points. And 10th to 90th is 82.4 to 85.6, a total of 3.2 percentage points with imputation in approximately, as I understood it, 20 percent of the values. That looks like an incredibly narrow range to me.

DR. SCHMITZ: It is a narrow

range. In part, this compression is due to the way in which the scores were scaled at the very end. I must admit when we started this process we stopped after the initial scaling, and there were scores that went from, because they're normalized they went from minus something to some other number. And we encountered enormous resistance because, for us, having a negative number didn't mean anything in particular. But, obviously, for a hospital, having a negative score would be a terrible thing.

So these scores were scaled in a particular way, and they were scaled so that zero represented the worst score you could possibly get. You would have to have zero for everything. None of your patients survive 30 days, all of your patients were readmitted, you didn't do any of the process measures. And 100 represented the best possible score. We did that because those were the upper and lower bounds that were

possible to define in a natural way. But no hospitals got anywhere close to being near zero because nobody is doing, nobody has those things happen.

The result, though, of that process of scaling to ensure that all the scores would be positive meant that the reported scores would be compressed in this way. There's another way of scaling that would spread them out more. And I should emphasize that those scores that appear in the right two columns are not percents, they're scores. But the compression of them does, in large part, is a result of the strategy we used for scaling.

CO-CHAIR GIBBONS: As well as imputation at the national --

DR. SCHMITZ: Yes, yes, that's right.

CO-CHAIR GIBBONS: -- because the outcome measures are drivers of a fair bit of this spread, and they're imputed in many

cases.

DR. SCHMITZ: In many cases. Right.

CO-CHAIR GIBBONS: All right. So I would like to suggest that we now vote on scientific acceptability. Partially nine; seven minimally; five not at all. All right. We'll now move on to usability.

DR. THOMAS: So this measure is currently not in use. CMS proposes a dry run of public reporting in the second quarter of 2011 if this is endorsed by NQF. And there is harmonization within each domain of the process of care and outcome of care domain, and, in theory, this is, of course, additive value if we could have a composite measure for acute MI. It would be very useful to consumers. Again, I do have some concerns about the usefulness to providers.

DR. AYALA: Can I ask a question? Because the gentleman who speaks about the

mathematical aspects of this keeps referring to the fact that it's created basically for the providers to use to assess their performance, but we're also talking about it being publically reported. So our concerns about the accuracy or the scientific basis for the composite I think becomes even more concerning when you're talking about publically reporting.

DR. SCHMITZ: Actually, we have argued really that the composite is aimed primarily at consumers rather than at providers. Most of the providers we've talked to have emphasized that, from their perspective, the individual indicators, the individual outcome and process indicators are the vehicles by which they gauge their performance and it is only by improving on those that they would increase the value of their composite. So we're not really arguing that this is of primary use for providers.

DR. RICH: Our experience in Detroit with public reporting is that while we had hoped that it would really be utilized by consumers it's actually much more highly utilized by providers and has improved care by providers looking at it and wanting to improve. So we don't have, we've tried very hard over the last four years to strongly get the consumers engaged, but I think the greater utilizers are the providers.

CO-CHAIR GIBBONS: I think we should go ahead and vote on usability. Quite a spread. One completely, nine partially, eight minimally, and three not at all. Feasibility.

DR. THOMAS: The data is generated through coding, and there is electronic means to acquire the data.

CO-CHAIR GIBBONS: So all these elements are available. Any other comments on that? So feasibility. So seven

completely, ten partially, one minimally, and two not at all. So then the final vote for endorsement, does the measure meet all the NQF criteria for endorsement? Seven yes and fourteen no. So we're going to move on now to 282, angina without procedure. And Roger Snow is our primary reviewer. Roger?

DR. SNOW: All right. This is a bit different kind of measure. This is a prevention quality indicator which uses hospital data to inform us about something else. The background is that in 1993 the Institute of Medicine published a monograph in which they called attention to ambulatory care sensitive hospitalization. The issue was that there are a lot of quality issues out in the community that have been very hard to measure that might be measured by looking at hospitalization. The argument behind that being that there is a series of conditions that if you're getting good access to good care you won't have to go to

the hospital, so that if you are discharged from the hospital with those conditions it argues that you weren't getting that kind of care before. That's the concept. And they argued for this, and AHRQ developed a total of 14 preventive quality indicators, the PQIs, one of which was angina without procedure. The metric is discharge from the hospital where the discharge diagnosis of angina without having had any of a long list of procedures which would include things like PCI and stenting and heart valves and the list is really remarkable. The argument for that being that if you were discharged with that diagnosis but didn't have a procedure, well, you probably had chronic stable angina. Ray has written about that, and it raises the question again of your access to or quality of your ambulatory care. That's the concept.

The issue has been around for a while. It's been adopted by several states,

it's been in use, we will come to that. But there were some problems, and I'm not quite sure, maybe we should bring it right up now. Very early on, there was a paper published in which they raised the question that -- this is the paper in "Health Affairs" -- socioeconomic status accounted for a lot of these hospitalizations and that that needed to be somehow embedded in the measure. The measure does have factors for age and gender.

And then there was a subsequent paper published titled "No Pain but No Gain" in which the authors noted a sharp decrease in the number of cases where there was a discharge diagnosis of angina. They then dug into that using the control numbers of the SEER Cancer Registry, which is a nationally recognized public registry. And what they found was that there was a reciprocal increase in the discharge diagnoses for coronary atherosclerosis.

They then looked at the incidents of AMI and that was unchanged. They then looked at what happens with people who were admitted with a diagnosis of angina and were they discharged with a diagnosis of angina or coronary atherosclerosis, and what they found was, although it's not quite as dramatic in appearance, the same crossover, that there was a significant increase in coronary atherosclerosis diagnoses and a sharp decrease in the angina diagnoses. Well, this makes a real problem because that decrease would normally, as the measure was intended, indicate one of two things, either that everybody was getting a procedure, because that was an exclusion phenomena, or that the care had dramatically improved. And the other data just didn't include that, so their conclusion was this decrease in angina hospitalization discharges was merely due to a change in practice in how people coded the darn thing. So that raises the

whole question of the viability of the measure. What started out as a really serious attempt to use this innovative and interesting concept fell apart in this case in the opinion of these authors, and I'll say that I was quite persuaded by that argument.

So there were these two separate problems with interpreting the data. And if you're not measuring what you think you're measuring, then you're not measuring anything at all. And I came to the conclusion at the front end that this probably should, even though I think it's been previously endorsed, with the development of this additional information, should not be considered on the basis of importance simply because you can't say something is important if you can't reliably measure it.

CO-CHAIR GIBBONS: Are the measure developers on the phone?

MR. BOTT: Yes. This is John Bott with AHRQ, and I think I'm joined by a couple of others if they'd like to introduce themselves.

MS. DAVIES: Yes, this is Sheryl Davies from Stanford.

DR. ROMANO: And this is Patrick Romano from UC-Davis.

DR. GEPPERT: Jeffrey Geppert from Battelle.

CO-CHAIR GIBBONS: Do you want to comment in response to Dr. Snow's concerns?

MR. BOTT: This is John. I defer to one of the other folks on the teams, if they'd like to make any comment.

MS. DAVIES: This is Sheryl Davies from Stanford. Yes, his summary is accurate. The measure was endorsed. The study that he referred to has been published and certainly points to a decrease in coding for this procedure -- I am sorry for this condition -- and for a coding for angina to

coding for CAD. And regarding the SES, we do have optional socioeconomic data risk adjustment, an indicator for that may be a moot point for the change in the coding.

CO-CHAIR GIBBONS: Can I ask you to comment on a statement made in the last paragraph on page three, which says, "This indicator has unclear construct validity because it has not been validated except as part of a set of indicators?"

MS. DAVIES: Sure. This refers to the fact that most of the prevention quality indicators, when they've been studied in the literature, they've been looked at as a set. So all of the information that we have about their relationship with measures to access, the relationship with measures of access to care or proxies of access to care, socioeconomic status, are based on the relationship with the prevention quality indicators as a whole or similar set as a whole. There's little

information looking at angina by itself and its relationship to access to care.

DR. ROMANO: This is Patrick Romano. I would just add there are some exceptions to that. So, for example, there is robust literature looking at heart failure separately, looking at asthma separately, looking at diabetes separately. But the less common PQIs generally have not been looked at separately in a research context to establish a construct validity.

CO-CHAIR GIBBONS: Okay. That's very helpful. Are there other questions or comments from the committee? David?

DR. MAGID: Yes. I have two comments. One is with regard to the diagnosis of angina in the setting of troponin. So we know that, over time, that the proportion of patients with this diagnosis has gone down for a couple of reasons. One is there has been a decline in coronary artery disease. But separate from

that, there have been dramatic changes in coding practice with the advent of more sensitive biomarkers for acute MI. I also think that, unlike primary hospital discharge diagnosis of 410, which has a pretty high positive predictive value, that angina is not considered to be a hospital discharge diagnosis that has good performance characteristics. There's quite a bit of variability and, in fact, it's often hospital coders, not clinicians, who actually assign this diagnosis. So that's the first concern I have, and I think that's a major concern, just to be clear. And the second concern I have is this assumption that patients discharged in the hospital somehow should be getting procedures.

So I think we're trying to live in an era of medical care in which we are good stewards of resources and that there is a general feeling, if you look at the data that's come out of the folks from Dartmouth,

that there's wide variation in the use of procedures, specifically cardiovascular procedures, and that all the studies that we know of to date show that increasing use of procedures is not associated with better outcomes in this population.

So I'm a little bit concerned that this could tend to drive practice in the wrong direction in terms of cost effective care. Whereas, I understand how asthma admissions clearly are a sensitive indicator of perhaps poor primary care, the idea that discharges from the hospital for angina without an associated procedure is poor quality of care or indicates poor quality of care but somehow hospital discharge diagnosis associated with a procedure would not be counted in that way. It doesn't make any sense to me. So I think on those two grounds I would strongly say that we do not endorse this measure.

CO-CHAIR GIBBONS: Dana?

DR. KING: Do we have a measure that I've forgotten about or coming up of just the number of admissions for MI as an indicator of the quality of outpatient care? Do we have that? Does that exist as an outcome?

CO-CHAIR GIBBONS: The measure developers can comment. At least I'm unaware of that. That's one of the sort of fundamental issues in our healthcare system. If we do a good job, we should actually keep the patient out of the hospital. That's not measured. That's not measured. Did the measure developers want to comment on that particular issue?

DR. ROMANO: This is Patrick Romano. I would say that there's no existing measure that treats hospital admission for acute myocardial infarction as a bad outcome of the healthcare system. But one could certainly argue for such a measure, and perhaps it will be specified as

a measure in the future. With respect to the other concern that was raised, I would just comment on one point which is that the diagnosis codes in hospitals are always assigned by coders but they're assigned based, rather strictly, on physician documentation. So the underlying variation that we expose is primarily a variation in physician documentation. But having said that, I think that the point is very well taken that there has been a change in physician practice, and so it is quite unusual now for patients to be admitted to the hospital and discharged with a diagnosis of angina because usually the biomarkers are available in the emergency department and a specific diagnosis is established before the patient is actually admitted to the hospital. So there is that change in practice. And we've seen more recently, looking beyond the period of Saver's article, a further two-thirds decrease in

the rate of this indicator since 1999.

CO-CHAIR GIBBONS: Tom?

DR. KOTTKE: Yes. I'd like to, first of all, express my appreciation of the importance and the positive intent of trying to keep doctors from simply parking patients in the hospital while they try to figure out what to do with them. I'm concerned with the shifting diagnosis because I think we're all, I don't know if they do it in Montana but certainly in Minnesota we are instructed on coding, how to code and like don't use this, use that. And so this is very susceptible to shifts in coding. And like Dave Magid said, I'm concerned, it's probably not happening but driving doctors to do procedures. I mean, it's just much easier to change the code. But this is a very large group of patients that are admitted to the hospital, have negative biomarkers, and they also contribute a large proportion of post-hospital deaths in the

subsequent year. They're admitted because, in clinical parlance they smell bad. You don't know what's wrong with them. They're not having an acute infarct. There's a lot of these. They contribute a lot of deaths in the subsequent year, so there needs to be a lot of work to be done. But I don't think this measure is --

CO-CHAIR GIBBONS: But as David pointed out, though, the reliability of some of the ICD codes is - I think we've discussed this enough. Yes, Helen?

DR. BURSTIN: I just want to make sure people realize this is at the geographic level of analysis. It's at a population level, not the hospital level, not the physician level. So it's intended to be a community indicator of that, so it's a little different measure than we're used to looking at.

CO-CHAIR GIBBONS: Thank you for that clarification. So we're going to vote

on importance of the measure. It is a unanimous no vote, so we have concluded this measure. We're going to take a break, but before we take a break the Chair needs a little poll so that we're clear on how we're going to proceed for scheduling the rest of the day.

So, first, this deals with when people's plane flights are. So how many people anticipate that they will have to leave before 3:30? Okay. And how many of those will have to leave before 2:30? Okay. I need to talk to you three at the break, please. Okay. We will break, and we will try to make it 15 minutes, please. Thank you, everybody.

(Whereupon, the foregoing matter went off the record at 10:23 a.m. and resumed at 10:38 a.m.)

CO-CHAIR GIBBONS: So we've reviewed the plane flight situation. We will be tight, but I am reasonably hopeful

that we can achieve the goal if we move through the measures reasonably expeditiously and if we work through lunch, so there will be a working lunch today. And we will do some on-the-fly last-minute adjustments if we need to. I think that the discussion of the last measure demonstrated that it would be helpful emphasizing the importance of having a brief introductory statement by the developer and that was my error in not doing that the last time. So for this measure, 355, I think we have some AHRQ representatives on the phone. If they could comment in short, three to five minutes, specifically on the intent of this measure, 355, the bilateral cardiac cath rate.

MR. BOTT: This is John Bott. I'll make a statement of less than one minute, and if others want to jump in, Patrick or Sheryl go. This is, in this case, a hospital-level measure where the

previous one was an area-level measure. In this case, we're looking at the rate of bilateral cardiac cath and in those people who had a cardiac cath in the hospital. Again, this is using an electronic inpatient claims to calculate the measure. I'll let Patrick, Jeff, or Sheryl add anything they think is necessary in here.

DR. GEPPERT: Well, this is Jeff. I just wanted to point out the most recent major modification to this measure, which was that the measure underwent our clinical panel review process and one of the recommendations of the clinical panel was that we add to the list of indications, procedure indications. So that was implemented in Version 4.0 of RQI software.

DR. ROMANO: And I'm sorry. This is Patrick Romano. I was on mute. I would just contextually sort of clarify that I think this indicator is principally viewed as an indicator of overuse or potentially

unnecessary procedure or a component of a procedure when it is performed without appropriate indications. So as time has passed on, we have revised those indications based on input from clinical experts.

CO-CHAIR GIBBONS: Thank you. Bruce?

DR. KOPLAN: Okay. So this is number 0355. The measure title is bilateral cardiac catheterization rate, and the brief description is that the developer wants to look at the percent of discharges of patients with heart catheterizations in any procedure field who had simultaneous right and left heart catheterization, so how often was the right heart cath also done in addition. This is an outcome measure, as was mentioned.

So part one is importance, and there's some interesting data that is provided stating that there appears to be high levels of use of right and left heart

cath and there's a significant amount of regional variability. They quote some data from the mid 1990s that reported between 11 percent and 50 percent bilateral catheterization rates, and I'm ashamed to say that Massachusetts had a 48-percent rate of bilateral catheterization rates, and that was during my fellowship training, so I learned at some people's expense, I think.

But one thing I will mention, not to date myself in any way, but getting back to the point that it is interesting that this data that's from the mid 1990s is rather impressive and almost kind of astounding, but I would wonder about more recent data. It does seem as if later in the submission, because we have the developer on the phone I would ask them this question that it seems like you report in Section 2F and 2H the implication is that the rates are much, much lower now, if I'm reading that correctly. So is that correct

that I'm interpreting a less than two-percent bilateral catheterization rate from more recent data? Is that true?

DR. ROMANO: Well, partly, that's a result of a general downward trend, which has occurred over the last ten years, and AHRQ reports data from the nationwide inpatient sample on the HCAP web site that can demonstrate that downward trend. But, in part, that's also due to the specification change that we mentioned in the introduction, the inclusion of additional exclusions for indication, the effect of reducing the rate from what was reported in those earlier studies.

DR. KOPLAN: Right. And I guess some of this we'll talk about under the scientific part, but in terms of importance, because some of the earlier themes of the day have been, you know, how much bang for the buck do we get if something is a low incidence rate, I just wanted to make sure

that we kind of mention that under the importance section. But, nonetheless, I could -- and, also, in talking with some of my colleagues around the room during the break, it seems like, anecdotally, that people seem to notice a much lower bilateral catheterization rate in their hospitals and programs than the 10 to 50 percent that is mentioned before.

But, nonetheless, despite this, it's also noted in your report that the cath rate - the bilateral cath rate has been used as a quality indicator for Medicare data and rightfully so. And despite this significant downward trend that's already occurred in percentages, I would still think that this seems to be an important indicator.

CO-CHAIR GIBBONS: Other comments?

DR. PHILIPPIDES: Was this also reported at a regional level or the hospital level? Because they mentioned, as AHRQ

mentioned before, it was regional.

CO-CHAIR GIBBONS: I think we were told this was a hospital measure by the developer at the start; is that correct?

DR. ROMANO: Yes, that's correct.

CO-CHAIR GIBBONS: All right. Thank you. Other questions? Sid?

DR. SMITH: Yes. When we start saying, I want to be careful that the message we're sending is not the fewer right heart catheterization you do the better you are. That is, an ideal hospital would do none. What I am seeing, first of all, is a relatively low rate, but there can be a tendency, as someone who is staffing and working in a cath lab, to rush through the right heart cath. The hemodynamics may not be done carefully or even patients with congenital heart disease or valvular heart disease where the information derived from a right heart cath with careful attention to left ventricular hemodynamics would be very

important in terms of decisions about surgery and management post-operatively is not done or it can be done hurriedly. Particularly, having worked both in the academic and private community, there can be a major focus on coronary anatomy to the exclusion of everything else.

So it seems to me the ideal and, you know, they say that man's reach always exceeds his grasp so maybe I'm asking for too much, but the ideal would not be to reduce catheterization bilaterally to zero but to recognize those hospitals where it is done only in very appropriate circumstances, that is complying with the guidelines and the evidence we have where it is of value. Have we done that here?

CO-CHAIR GIBBONS: Bruce?

DR. KOPLAN: Actually, your point is very well made because certainly, in the very beginning, the way the description is is that they're just looking at the percent.

And so to slightly paraphrase what you said, it seems as if you'd rather, you would like a measure that looked at a way of looking at the percentage of inappropriate or non-indicated right heart cath as a measure. And I think when we get to the scientific part two, the way the developer develops the numerator and denominator, it actually does express that. So I think the title is a little misleading.

CO-CHAIR GIBBONS: Okay. I think we should vote on importance. So the vote is 18 yes and 3 no. We'll move on to scientific acceptability. Bruce?

DR. KOPLAN: Okay. So this is where it gets a little bit interesting and a little bit of a -- I have a few questions here. So the developer, just carrying on with what we were just talking about, defines the numerator as discharges that have coding for right and left heart cath, and they exclude, it seems as if there's a

long list of exclusions that would exclude diagnoses that would lead to an indication for right heart cath. So what it seems as if they're trying to do is eliminate, is to only count what would be perceived as non-indicated or inappropriate right heart cath. So that was my take on the numerator.

The denominator looks at heart catheterizations in any procedure field but only to include cases with coronary disease. And I wanted to step back and think about that a little bit because should any non-coronary artery disease type cases be included in the denominator, I don't want to be nit-picky but I just wanted to ask the developer or ask the group if they felt that this type of numerator and denominator actually expressed what they were trying to achieve. That was all. So that's as far as numerator and denominator goes.

In terms of, it seems as if reliability and validity testing have been

done using large databases, and I think that's all I have to say about the scientific aspect, if anyone has any comment.

DR. RICH: I was just looking at the disparities information in this section, and it's kind of interesting we're looking at it by payer. Sorry, I scroll up and down it is around 2 -

CO-CHAIR GIBBONS: 2H?

DR. RICH: Is it 2H? Thanks.

CO-CHAIR GIBBONS: Yes, 2H.

DR. RICH: Okay. So what I'm seeing is that you actually have less, like the rate goes down for, it's highest for Medicare and lower for Medicaid and lower for other. I was kind of curious about that. It's a little bit, to me, perhaps counterintuitive but I'm not really sure, but I was wondering if the developers had any comment on that.

CO-CHAIR GIBBONS: Did the

developers hear that question? It's about Section 2H of the application where you show different rates for Medicare, Medicaid, and other payers. The question is really are they real and have you done anymore analysis on those?

DR. MAGID: If they're not age adjusted, you'd expect the Medicare rate to be quite a bit higher.

CO-CHAIR GIBBONS: Okay. So I guess we have a follow-up question. Are they age adjusted?

DR. GEPPERT: No. The indicator itself is not risk adjusted.

MS. DAVIES: But the strata that are reported out --

DR. GEPPERT: Strata, yes, yes.

CO-CHAIR GIBBONS: So do the difference across payers just reflect difference across ages?

DR. GEPPERT: Probably.

DR. KOPLAN: And then while we

have the developer, the question I asked, can you express why you just did CAD patients in the denominator? I guess that encompasses pretty much everybody, but should anyone else be included in that denominator?

MS. DAVIES: So this indicator underwent a clinical panel review, as you see in the documentation. And their recommendation was to stay with the CAD in the denominator, you know, and taking out the indications that are exclusions in order to really hone in on those patients for which right heart catheterization is most appropriate.

CO-CHAIR GIBBONS: Yes. And, actually, after thinking about it, I thought that that made sense because any non-CAD patient would probably, as you've said, would be excluded anyways from the numerator. So I think that, I think I'm okay with that.

DR. KOTTKE: If you reflect on cath lab burden or, you know, the non-CAD is such a small part that trying to figure out exactly who in that non-CAD ought to be in the numerator and denominator probably is burdensome.

DR. PHILIPPIDES: Bruce, you cited some data from your era when you were a fellow, skewing all the numbers. But it appears that in 2F there might be what I'm assuming are more soft of recent numbers, getting 5th, 25th, median, and then 90th percentile results using, I think, the new exclusion criteria. And correct me if I'm wrong, is it varied between 1 percent and about 2.4 percent?

DR. KOPLAN: Yes, that's what I kind of alluded to at first is that these data from the 90s are up to 50 percent but more recently the rate of right heart cath, my take on that is that the rate of right heart cath in people undergoing left heart

cath is very low.

DR. PHILIPPIDES: Right. Now that they've added the new exclusions that take into account the appropriate right heart caths which we think are important, it seems like the discerning capability or the gap is really narrowed quite a bit. So it raises the question as to whether or not we're going to be able to discern as much as initially implied.

CO-CHAIR GIBBONS: Right. So George has pointed this out for those who don't have the numbers in front of them, 5th percentile is 0.011, 95th percentile is 0.0246. So the 5th to 95th is a difference of 0.013, in other words a 1.3 percent difference in the rate of inappropriate right heart caths, as stated here. Is that correct, developer?

DR. ROMANO: That strikes me as a little bit narrow, but that's approximately correct. A little more recent results we've

been showing have a variation from about 1 to 4 percent from the 5th to 95th, but same order of magnitude.

CO-CHAIR GIBBONS: And this is a national number, and you don't have regional variation numbers of this form? Because it seemed like the older data expressed a lot of regional variation, and if you showed these numbers but then followed by a tremendous amount of regional variation it would give more weight to the issue, I guess. But we don't have that.

DR. ROMANO: We don't have regional data for the new specification. Under the old specification, the most recent data was 2007. It shows, you know, rates of about 3 percent in the South and the Midwest and about 5 to 6 percent in the West and East Coast. So a doubling of the rate across regions.

DR. KOPLAN: So is this a renewal then? It appears that this is working and

maybe it raises the question of how much more can be done with it. Just out of interest, I notice that the denominator or the sample comes from 4,000 hospitals, and I remember, I think relatively accurately, that we only had 2,200 hospitals reporting on the use of ACE inhibitors. It's remarkable to me that 4,000 hospitals are doing right heart caths but only 2200 have echos or are able to assess LV function non-invasively. Is that 4,000 number correct?

DR. ROMANO: It refers to the total database that AHRQ maintains but for the state and patient databases. The actual number of hospitals in the bilateral cath indicator is a little over 1,900, so it's comparable to the 2,200 number.

CO-CHAIR GIBBONS: All right. So I think you've relieved Dr. Smith's anxiety.

DR. SMITH: Yes. I was looking at 2F.1, but it sounds like that's a larger sample, that the cath labs are much closer

to 3,000?

CO-CHAIR GIBBONS: 1,900.

DR. SMITH: 1,900. Okay, good.

CO-CHAIR GIBBONS: All right. I think we'll vote on scientific acceptability.

DR. RUSSO: I think this does overall, though, raise the concern of selecting measures that we see that we might perform better on as part of that possibility and that really may promote the use of more composite measures. So we've eliminated beta blockers because such a small number were in that group, but composite measures may be the way to eliminate any selection bias.

CO-CHAIR GIBBONS: So ten completely, nine partially, two minimally. Let's move on to usability. Bruce?

DR. KOPLAN: Yes. In terms of usability, the measures appear to be in use in multiple state and some national

reporting agencies, and so it seems like they've been demonstrated to be usable and there do not appear to be any particularly harmonization issues that I could see with this measure.

CO-CHAIR GIBBONS: Other comments? If not, let's proceed to vote on usability. Fifteen completely, five partially, and one not at all. And now feasibility.

DR. KOPLAN: And, once again, the data is generated from coding, which should be easily obtained and should be able to be obtained from electronic record sources. I didn't have any major issues with the feasibility part of things.

CO-CHAIR GIBBONS: All right. And we'll vote on that now. Seventeen completely, four partially. And then, finally, we need to vote on endorsement of this measure. Seventeen yes, three no. So this measure is approved for endorsement.

We're going to move on now to measure 133. Before we move on to this next set of measures, though, we're going to hear from the measure developer, the ACCF. And demonstrating that he wears many different hats, that's Dr. Masoudi.

DR. MASOUDI: Good morning. I'm Fred Masoudi. I'm here as the senior medical officer of the National Cardiovascular Data Registry. The next four measures that you're going to be looking at are those that have been submitted by the NCDR, which is a joint effort of the ACC Foundation and the SCAI. The registry itself collects data on patients undergoing catheterization and percutaneous coronary intervention in approximately 1,100 hospitals, which represents about 70 percent of the hospitals that perform PCI. It includes about 80 percent of patients who get PCI nationwide. I won't go through the importance of PCI other than to say it's

probably one of the most widely performed invasive procedures in patients with cardiac disease and is associated with substantial expense.

The three measures that you will look at, one is an outcomes measure that has already been endorsed and so is up for reassessment. It is a risk-adjusted mortality model. And I'm joined by Matt Roe who is one of the developers at DCRI on the phone, as well as ACC staff, to discuss that, as needed. There are also three process measures that you will look at, one of which is clopidogrel at discharge, one of which is aspirin at discharge, and the third of which is statins at discharge. I would say a few things about these. First of all, these are harmonized in terms of their specifications with the existing CMS measures that look at patients with acute myocardial infarction. This is a different denominator of patients. These are all

patients who undergo PCI, only about 30 percent of which have acute coronary syndrome. The remaining 70 percent are receiving PCI for elective reasons.

As you'll notice, the performance data for the clopidogrel and aspirin are high. You know, you've had prior discussions about some of the other measures. The statin performance is markedly more variable. A couple of issues about this. One is that, again, although this data is collected from the majority of hospitals that perform PCI, because this is a voluntary registry, this isn't necessarily nationally representative. As Dr. Smith pointed out, performance typically tends to be lower in sites that don't voluntarily participate in quality programs like registries.

The other issue is one that we can get back to later, which is this issue of the extent to which measures that aren't

endorsed can be included in composites, as the ultimate goal of the measurement program is to generate a composite measure for the use in public reporting. So that will become relevant later on, as well, on during these discussions. Thank you.

CO-CHAIR GIBBONS: Thank you, Fred. So the primary reviewer of measure 133, PCI mortality, is Sid Smith.

DR. SMITH: Thank you. And I think you've heard a good description. This is a renewal of a program underway which has been very productive. The risk-adjusted PCI mortality is an outcome measure, and I think you've also heard this is a very frequently performed procedure which can have a major impact on patients' lives to the better if it's done well and very detrimentally when complicated by death. It's very expensive, and so some information and knowledge about how our centers are performing and how we can do better is really very important.

It's a frequently performed procedure.

The mortality that we're talking about here involves all PCI procedures performed clinically, that is both for acute coronary syndromes and for chronic coronary disease. So it's not separating out there. And I think it's been underway. It has been very productive. There's some really nice publications from the existing registry, one that involved over half a million patients in JACC recently, Journal of the American College of Cardiology.

I have some thoughts about the science which would help me in terms of reassurance. But I think this is a very worthy program, so I guess we need to vote on the importance of it.

CO-CHAIR GIBBONS: Are there other questions or comments before we vote on importance? All right. Let's go ahead and vote on importance. Okay. The vote is unanimous, 21 yeses. Sid, scientific

acceptability?

DR. SMITH: Okay. The science here and some of my concerns may just be the problems that are inherent in all registries, but it's the exclusion criteria. I guess the first thing is are all patients reported consecutively, or is it only those patients that the operator sits down and fills out a report or the hospital reports him? Is there a way to be reassured that those hospitals participating in the registry are actually reporting all patients? If they aren't, then it's highly possible that patients where there are major complications, who die, might not be reported, and that could give us an underestimation of the true percentage of mortality. So are consecutive patients reported? Is that a requirement of the registry?

And the second thing that concerned me a little bit was that the data

submissions that don't pass a data quality and completeness assessment are apparently excluded. So, again, a sloppy report doesn't enter into the assessment, and it would seem that eliminating sloppiness is one of the things we want to do here. So we need to know about the data as a whole, so I'm a little concerned about the effect that excluding reports because of completeness might also aim the - or bias the mortality to be lower than it actually could be.

CO-CHAIR GIBBONS: Okay. We'll refer both of those questions to --

DR. SMITH: And then there's a third one.

CO-CHAIR GIBBONS: Oh, sorry.

DR. SMITH: The third one is that, and I've sort of come up with this in other registries, but it's a decision to exclude patients from this consideration if they have more than one angioplasty with an admission. What bothers me there is that a

patient who comes in who may have had a stent delivered inappropriately goes back for an operation or for a procedure that's related to poor performance of the first and, because of that, is excluded and dies, so there's no way to really, you lose that population of patients. Do you understand what I'm saying? Okay. So those are the three things.

CO-CHAIR GIBBONS: Three questions. Okay.

DR. SMITH: Consecutive data, the exclusion of improperly filled out reports, and the idea of how we handle those patients who may have been taken back for a second procedure that was actually an urgent procedure related to a poorly performed first procedure.

CO-CHAIR GIBBONS: George?

DR. PHILIPPIDES: The people who get involved in registries might have addressed this, but the idea that patients

that are transferred to another facility are also excluded. It seems to me, oftentimes, the sickest of the sick or the people that aren't doing well for reasons that are not always captured in these kind of registries get sent out to, it feels like to my hospital when I'm on call. Yet, when they come -- sorry. Mark will take it up from here. All kidding side, and then when they come to us, there's no place else to send them, so those patients stay with us and the mortality becomes part of --

DR. SMITH: Yes, you get hit with the mortality.

DR. PHILIPPIDES: In this measure, but there are some other measures where they don't take out the transfer issue. So I just wanted to raise that, as well, that the high-risk patients are being sent out.

CO-CHAIR GIBBONS: All right. I think we now have four questions for the

developers.

MS. FITZGERALD: Okay. So the first question is --

CO-CHAIR GIBBONS: Please identify yourself.

MS. FITZGERALD: I'm Susan Fitzgerald. I'm one of the staff at ACC. Sorry. By contract, hospitals are supposed to submit all records. Now, that's by contract. We police that with our audit just to verify the count of records by a cath lab log or some other independent one. We've tried to do that other ways. It's been logistically difficult, but we do it with an audit. Last year, we audited 25 sites, and that's part of the audit. That's the first thing.

I don't know if Tony is on the phone, but our data quality program, it's in the very high 90s the percentage of hospitals that submit data that pass the data quality thresholds and are included and

get risk adjustment. Many of the hospitals, almost all of them, are in some form of pay-for-performance program, so to get their reports is important to them. If they're not included, we're not going to send their data to, we call them our analytic research service clients. So that's number two.

Second PCI, it might be misleading. I don't know if Dr. Roe wants to speak to it, but what we do is we look at there are variables during the PCI that are looked at. So the second PCI, the patient is not excluded, but the procedure is. So we're looking at a patient admission or a patient record not a procedure. So what we're saying is that we're looking at variables in the first procedure within the model but not in the second procedure, so the variables in the second are excluded.

And then the patients that are transferred to other hospitals are excluded because we don't know if they lived or died,

and we know they probably were sick and something else was happening, so they are excluded in the model.

DR. ROE: This is Matt Roe from the Duke Clinical Research Institute, just to follow up on the comment that Susan made. The model is focusing on the first PCI done for a patient during a given hospitalization, and I think there's a good point made before that if that procedure is performed inappropriately or there are problems and the patient had the second procedure and the model may not be accounting for that. In some sense, that's correct, but the patient's hospitalization will still count in the mortality. If that patients dies, it will still count. But it's really hard. You can't re-frame the model on a second procedure after you have already done it on the first procedure. I think that becomes very difficult, and it's also an infrequent phenomenon within the

registry. So it would be hard to even develop a model that could do such an aspect there.

And then the transfer out part, I think recognized this is an inpatient mortality model, so if a hospital actually does PCI and then transfers a patient out to another center you cannot capture what happens to them after transfer, so we don't know whether they lived or died. So they have to, by nature, be excluded. But we recognize that that's a very infrequent phenomenon as well because when a patient gets transferred out after a PCI it may typically only be for a patient who needs urgent surgery at a center, for example, where the PCI center is doing PCI without on-site CABG facilities. So, again, I think those scenarios are pretty unlikely but difficult to really overcome them with the way the database is structured and the way the model was developed.

DR. SMITH: So I'm happy with the explanations. I think the area that, how we handle selective reporting or poor reporting of events would be something to go after in terms of the registry. Those hospitals that have poorly completed reports should be audited in some way in an effort to be sure that all reports are entered and entered correctly. But, again, I think the handling of the transfer out is appropriate. And from the earlier comments, I think they are handling -- the first procedure does get entered, and if the patient comes back because of a complication it's reflected on the first procedure. So my concerns there are handled well, I think.

Now, I just, again, have to say that the database here is robust and the observations have provided very important information about PCI. One thing I did not mention in my introduction was there still is a gap in terms of mortality after PCI if

you look at it among the different hospitals, and this database allows hospitals a comparable volume to compare themselves against each other and also against a national baseline. So, overall, I think that this project has been very valuable. But I think that there are miles to go before we sleep --

CO-CHAIR GIBBONS: Okay. So we need to move on. We're falling behind schedule now. We've got to speed up a little bit. Scientific acceptability? Okay. Completely 13, partially 7. We'll move on now to usability.

DR. SMITH: Well, I think however it's said, res ipsa loquitur, it speaks for itself. This thing has been used very well by many hospitals, so I think it's demonstrated that it can be done and it can be done with very large enrollment. The major issues that hospitals face is, I think, how to get their data, who's going to

enter the data, but it appears that that's been done well here.

CO-CHAIR GIBBONS: Okay. Any other comments or questions?

DR. SANZ: I'd just like to point out that the usability is based on the accuracy of the data, which I know is self evident. But if you have an institution with an outpatient cath lab and one of the exclusions is transfer to another facility, you will easily have, I shouldn't say easily but almost easily have a zero mortality by transferring a patient who has a complication with CPR to the local nearest bypass surgery facility. And this is an issue, and I think it also comes into some of these other measures where you have an exclusion with something that occurs very infrequently but transfers the mortality to another institution. And I don't agree that you can't find that data. It's a simple phone call to another institution. You know

where they got transferred.

CO-CHAIR GIBBONS: Dr. Masoudi?

DR. MASOUDI: Just to speak to that, so you wouldn't have a zero mortality if you don't submit any patients to the registry. It's only the patients you submit. So a site that does not participate that's an outpatient site, which generally wouldn't be submitted to the registry anyway because it's an outpatient site, simply is not captured in this data. And you could say it is true that this doesn't capture data on outpatient cath labs, but it does capture data on a large number of hospitals and is useful for hospital reporting. But just to clarify, a transfer out is excluded, so that would not count towards your denominator, for which you would have a zero in the numerator. Those patients are simply excluded.

CO-CHAIR GIBBONS: David?

DR. MAGID: So I think Sid

brought this up as an issue. I think it is a little bit of an issue, probably not a lot of an issue, in the sense that institutions that don't have complete capabilities may transfer out patients for CABG or for other major procedures. And, you know, Matt, this is a suggestion and you might consider looking at within the data set at the sites that don't transfer out and the characteristics of those people that transfer. You could probably build a propensity model that imputed mortality on those transfer people and run the analysis both the way you're doing it now and estimating mortality in those transferred out for the institutions that do it with some regularity just to see if there are any difference. So just a suggestion.

DR. ROE: That's a great suggestion, and we'll certainly take that under advisement. Again, I don't have data in front of me right now, but I think the

transfer out rate is pretty small, but I think in centers that do it more frequently it's a good idea to see how that comes out.

DR. STEARNS: My concern would be that for consumers looking at these numbers, if the patients that are getting transferred out, if the mortality is not included, aren't we giving them a false sense of security if they're looking at these numbers?

DR. MASOUDI: You know, again, the transfer out percentages are quite small. And, generally, the bottom line is that you really cannot reliably identify them, and part of this has to do with a lack of a national patient identifier. There's no reliable way to identify what happens to a patient after they've been transferred. You could say it's just a phone call. The fact of the matter is if there's anything that could be gamed it would be that. So because you can't reliably identify what

happens to a patient after their transferred, they can't be eligible for inclusion in the measure. And this may affect a small number of centers. That is correct. But the overall, the proportion of transfers as a group of the entire data set is quite small. We have the numbers, 0.7 percent.

CO-CHAIR GIBBONS: Okay. We have to move ahead. We're going to vote on usability. Eight completely, twelve partially. All right. Now we have to move ahead to feasibility.

DR. SMITH: I think some of my comments on feasibility I made in the usability. The data are available. They are retrievable. The major limitation is being sure that someone enters it accurately, and I think the size of this registry, if I'm not incorrect, looking at the JACC article now, over 500,000 patients suggests that it is doable by a number of

hospitals. So I would say it appears to be quite feasible.

CO-CHAIR GIBBONS: All right. We're going to vote on feasibility. Twenty-one completely, a unanimous vote. All right. We'll move ahead and vote on endorsement. Okay. A unanimous vote in favor of endorsement.

We're going to move on to 1495, but I would point out that for the next three we really have to keep to schedule. If at all possible, 15 minutes each because the more we fall behind the less time we will have for the rehab measures before our reviewers leave. So Mark?

DR. MAGID: Can I just ask one question? And I'll be quick. Dr. Masoudi brought up this issue that they're looking to put these measures into a composite and that some of them may have high rates but that if NQF, if we don't endorse each of them separately it's hard to make them into

a composite. Is that correct?

DR. BURSTIN: They have to each be evaluated for endorsement. You can make the argument that a non-endorsed measure should be part of a composite for balance or something like that, but it has to be evaluated but not necessarily endorsed.

DR. MASOUDI: So would that evaluation occur here then?

DR. BURSTIN: Yes, right. Exactly. But they could still be in the composite if you can make the argument they should live in a composite, even if they wouldn't be stand-alones.

DR. MAGID: And after we are done if we say, well, we like these individually but they have high rates, we'd rather see them in a composite, we can say that at the end or --

MS. PACE: Right. But one thing to keep in mind, I think part of what you saw in that composite that we talked about

earlier with the CMS, besides the missing and imputation issues, part of that lack of variability was because all those component measures had high rates. So, you know, do you accomplish anything is another question.

CO-CHAIR GIBBONS: We must move on. We must move on. I'm sorry. The Chair has got to start getting tough here, or we're not going to get done. Mark, you're on.

DR. SANZ: Okay. Measure 1495 is Plavix post PCI at the time of discharge. This is like one of those ten commandments. It's just something that every interventional cardiologist knows and does. I will quote 1C14, which is this guideline is the most widely-recognized professional guideline in the U.S. for cardiovascular medicine in the area of PCI care, so importance is not in doubt.

I will say I believe that this should be a composite score with the next

one, and much of the comments I'm going to make relate to both. I understand -- side discussion with Fred that we have to somehow vote on these separately, but, in the end, they should be combined because nobody should go home in the United States off of both aspirin and a P2Y12 inhibitor. So in my opinion, it does not make sense to look at them as independent, but if we have to do that to get them together that's fine.

All I'll say is a lot of the data that's quoted really doesn't pertain to PCI. It doesn't change the value of the measure, but there's a lot of stuff in here about unstable angina, STEMI, et cetera, which has little to do with this measure, and it's repeated over and over. But the measure itself stands.

CO-CHAIR GIBBONS: Other comments vis-a-vis importance? Dana?

DR. SANZ: Is this where we talk about the gap, the performance gap?

CO-CHAIR GIBBONS: Yes.

DR. SANZ: Well, I'll just say the performance gap is extremely low, but I was going to talk a little bit about that under the next one. I will say that this concept that exclusions are not a problem, I disagree with. The gap here is trivial, 98 percent or so, and everybody is about the same. Of that, I would estimate personally, based on my institution and the local area, half or more is administrative miss. When you go back and look at each patient, they actually went home on it; it just didn't get documented appropriately.

Finally, to say that transfer to other hospitals isn't important, I did. Thanks, Tom. Back of the napkin calculation. If 0.5 percent of the one million PCI are excluded, which is what is in the data of the Plavix one, 0.57 I believe it was, which was the second largest exclusion, and half of those will get

subacute thrombosis, that's a pretty big number. You end up with about 1,500 deaths or infarcts. So I think that these exclusions, while small, if you look at the benefit to be gained, which is only one percent, if you keep measuring this, it's about the same.

So this may be one of those things where the performance gap gets low and then why not just get rid of the exclusions. Obviously, death has to remain. If you're doing this and the patient ends up in hospice, there are some issues as to what you're doing. Maybe you didn't know they were going to go to hospice, but I question that all of those are appropriate. Maybe it's because of what you did that they ended up in hospice.

CO-CHAIR GIBBONS: Developer? Dr. Masoudi?

DR. MASOUDI: Yes. Just in terms of these exclusions, again, the attempt here

was to harmonize these measures to the extent possible with existing CMS measures, which use these very selfsame exclusions. And so in order to minimize burden on practitioners, these are specified identically to the CMS measures. I guess the argument that was made about, you know, a 0.5 percent miss in exclusions could have this immense impact is more of an argument to accept this measure even though performance rates are high because even marginal increases then in improvement would lead to substantial improvement in health outcomes.

But I think the key factor is in terms of the exclusions these are completely concordant with those that are used for CMS for all of its inpatient measures. That was done intentionally as a means of trying to reduce provider burden.

CO-CHAIR GIBBONS: Okay. We need to vote on importance. Okay. Unanimous

vote in favor of importance. Let's move on to scientific acceptability.

DR. SANZ: I have one concern under 2A.3, which I suspect is only an English problem, but it says that it should not be there if PCI is attempted. Frankly, if PCI is done and there's no stent placed, you don't need Plavix. So this should be with a successful stent implantation and remove angioplasty itself. I think it's just an English language thing that needs to be changed because right now we're going to include angioplasty without a stent. Does that make sense?

Under 2A.4, it's actually stated correctly. I'll just stop there. I think that, otherwise, it's very good.

CO-CHAIR GIBBONS: So does the developer accept that correction?

DR. MASOUDI: Yes, that's intended that way. We accept the correction. I'll put it that way.

CO-CHAIR GIBBONS: Thank you. All right. Any other comments or questions? So our vote is in light of that correction, accepting that correction. Nineteen completely, two partially. Moving on to usability, Mark.

DR. SANZ: It's being used, well, certainly everywhere the NCDR is, which is most places, and used well. Harmonization, there is another measure for drug-eluting stents separately, which I don't really understand. I don't think it necessarily came from ACC, but that needs to be harmonized. There's no reason to have both of these. And then the issue of dual antiplatelet therapy, there really should be one measure which somehow we need to vote on later.

CO-CHAIR GIBBONS: Okay. Additional discussion on usability? All right. Let's go ahead and vote on usability.

Seventeen completely, four partially. Now feasibility.

DR. SANZ: I don't have a whole lot to say here. I do think it's feasible to make a phone call regarding what happened to a patient transferred out. The N is small. The number of those that are going to have a major event is high, so that's it. Otherwise, it's already being done.

CO-CHAIR GIBBONS: Okay. Let's go ahead and vote on feasibility.

Seventeen completely, four partially. And then, finally, let's vote on does this measure meet criteria for endorsement.

All right. So that's a unanimous vote for endorsement.

We'll now move on to 1493, aspirin at discharge for PCI. Mark, you've already indicated you think the whole construct here is very similar.

DR. SANZ: It's nearly identical.

You change the drug, but as far as any -- I didn't see anything specific --

CO-CHAIR GIBBONS: David is about to weigh in with his usual motion.

DR. MAGID: Yes, I think dual antiplatelet therapy is noncontroversial, and I think I would suggest that we consider the same vote for clopidogrel that we did for aspirin.

DR. SANZ: And I want the record for the quickest.

CO-CHAIR GIBBONS: Does the developer see any difference in this measure versus the clopidogrel measure?

DR. MASOUDI: No. I think you could do a find and replace with aspirin and clopidogrel essentially. Again, the contraindications to aspirin may be somewhat different from those of clopidogrel because it would be an aspirin allergy and not a clopidogrel allergy. But, essentially, there's no difference.

CO-CHAIR GIBBONS: Okay. Are there any objections to simply duplicating our vote on this aspirin measure to be the same as we just voted on for clopidogrel?

DR. KOPLAN: So just to clarify, this is with the stipulation that it's intended --

CO-CHAIR GIBBONS: Well, we're going to have another vote on that issue. This is just to say for this measure, aspirin after PCI, we're going to vote the same as we just voted for clopidogrel. Is there anyone who objects to that? Okay. So Mark will get the record for the fastest turnaround.

So now I would propose that we have a separate vote, and that vote is to encourage the developer to combine these two as being aspirin and clopidogrel, i.e. dual antiplatelet therapy. And how do we vote on this? By hand I guess. All in favor of that? Opposed? All right. So I think we

will convey to the developer our formal suggestion that those two be combined.

All right. Now before lunch we need to do 1498, statins at discharge. Dana, you're the barrier from lunch. Don't feel any pressure.

DR. KING: The importance of this is not widely debated. Statin therapy reduces the risk of coronary events and coronary artery disease following PCI. This measure will encourage improvement in the rates of statin prescribing. Unlike some of the other measures we've discussed this morning, there is a performance gap. The prescribing rate actually from the 5th to the 98th percentile was from 72 to 98 percent. So there are people achieving the 98 percent rate, but there are a significant number of hospitals that are down at -- below that, and half of hospitals do not -- have over 10 percent of people discharged not on a statin.

Interestingly, they did do some stratified analysis, and the lower SES hospitals did as well or better than the big cats. So I just thought it was worth mentioning. So I guess I would not argue with importance of this measure, and there does appear to be a performance gap.

CO-CHAIR GIBBONS: Any comments or questions regarding importance? All right. Let's go ahead and vote on importance.

DR. KING: Okay. The scientific acceptability measure, the specifications are well done. It has the same reliability and validity as the ones we were talking about with Plavix and aspirin. It's taken from a registry, not from a total population thing. But, otherwise, it has the same reliability and validity as that registry, which was, as we heard, fairly complete and pretty reasonable.

As I mentioned, there was some

stratification being done, and I think it might be important to continue hearing about that since getting from the 72 to 98 might involve addressing some socioeconomic things, although right now it doesn't appear that they're different. And they did not do anything with other disparities besides SES, but I think the scientific acceptability is there.

DR. RUSSO: I have just a question. So on the exclusions for -- so I'm trying to figure out why there's such wide variability in that. Is it possible that some of the patients were -- it says contraindicated as obviously excluded, so an allergy. How about is patient refusal -- is that an acceptable -- or if they're placed on another agent for cost -- or is there some, I guess there's generic. Should there be any more -- is there any other way to figure out why there's such a difference across the country? It's not related to SES

so I guess it's not cost. Are there any other exclusions we're missing? Maybe not.

CO-CHAIR GIBBONS: Does the developer want to comment on that?

DR. MASOUDI: So basically there's a contraindication to statins regardless -- again, it's aligned with the CMS measure criteria where if there's a contraindicated noted that patient is excluded from the measure, unless the medication is given.

DR. SNOW: Is that the only exclusion? What about patients who opt for a different lipid-lowering agent? I mean, statins are wonderful and all but --

DR. MASOUDI: If the provider -- I think that happens in, I would imagine, a vanishingly small group of patients who didn't have an intolerance to a statin. But even if it did, if the provider didn't indicate that there was a contraindication to therapy, they would be indicated as

having failed. Again, I think that that is going to be extremely unlikely, but it's hard to know how many of those contraindications represent a situation where a patient who is tolerant to a statin has requested that their doctor not treat them with a statin.

CO-CHAIR GIBBONS: Any other comments or concerns about this? All right. Let's vote on scientific acceptability. Eighteen completely, three partially. We'll move on to usability. Dana?

DR. KING: Okay. This is currently in use and has been voluntarily reporting this measure. Participating institutions receive an institution outcomes report each quarter with their hospital's results.

CO-CHAIR GIBBONS: Any issues about usability? Okay. Let's go ahead and vote. Twenty completely, one partially. And now feasibility.

DR. KING: It's obviously feasible. It's being done. It's in use. The electronic sources are used, and even the survey is submitted electronically. They identified several paragraphs of their efforts to reduce inaccuracies and follow up on the process, and I think it was reasonable.

CO-CHAIR GIBBONS: Okay. We'll go ahead and vote on feasibility. One clicker isn't working. We have 20 votes saying completely.

All right. We'll move on now to the final vote. Does the measure meet criteria for endorsement? Unanimous support, 21 yeses. Okay. We're going to break for lunch, but we're going to have a working lunch, so I'd ask everybody to try to just grab lunch and get back in here, and we will restart on the next set of measures. Thank you to Dana and Mark for putting us back on time.

(Whereupon, the above-entitled matter went off the record at 11:47 a.m. and resumed at 12:02 p.m.)

CO-CHAIR GIBBONS: So this is a different block of measures and requires everybody to sort of, I think, listen carefully and adopt a somewhat different mind set as we approach these. So for that reason, they were originally allocated more time in the schedule, and we, hopefully, by getting back on time, will permit adequate discussion of these.

So we're going to open up by asking the measure developers to concisely, in three to five minutes, give us the intent of this block of measures.

DR. MARJORIE KING: Hello, I'm Marjorie King from AACVPR and the AHA/ACC AACVPR writing group for this measure. This group of measures was written when the referral to cardiac rehab measures were written and published back in 2007 to

accompany the measures. Those referral measures are NQF endorsed. These measures were written to set safety and performance standards for cardiac rehabilitation programs so that if we hold doctors to referral to cardiac rehab we want to hold programs to a minimum standard of quality.

In order to write these measures, we reviewed about more than 30 potential measures that would be appropriate. Not surprisingly, we ended up choosing measures that are very similar to the measures used in the AACVPR program certification process, and that's because the measures that we chose are based on the core components of cardiac rehabilitation and also on ACC and AHA guidelines.

You need to know about the AACVPR certification process because we used that process for some of our testing. First of all, it's been in existence since the late `90s. It has been an evolving process. It

is linked to standards. It is an all-or-none phenomenon. You either pass certification, or you do not pass certification, and you need to meet all of the specifications in order to be AACVPR certified. It was developed as a mentoring process, not as a pay for performance or anything like that process. And so when programs apply for AACVPR certification they have already attended seminars, had mentoring from their affiliates, and they don't apply unless they think they're going to pass. They wait until the next cycle, and they get all their ducks in a row.

The denial level is very low for AACVPR certification. It's about two percent. And when we looked at the denials for the last three years, the reasons for denials were across all four of the measures. They're very low numbers. We didn't put them in the application because they were such low numbers. But, again, you

have to understand the data that we had to analyze. Unfortunately, we do not know anything about the programs that are not AACVPR certified, which is probably at least 50 percent of the programs in the country. So we don't know what we don't know, and we don't know how to get about knowing the characteristics of those programs. Our measure testing, as I said, is very similar across all four measures because it used the AACVPR certification process for the inter-rater reliability testing, for example, and for the other testing.

We also ask ourselves the question, well, is there a relationship of AACVPR certification to what we're really trying to drive, which is improved patient outcomes using these processes that are stated in the measures. And so there is a large registry in Wisconsin with the Wisconsin cardiac rehab affiliate plus the Wisconsin Department of Health. We looked

at that data asking the question do certified programs have better patient outcomes compared to non-certified -- programs who are not AACVPR certified. That is in the appendix that is probably labeled Report to the Board of Directors of AACVPR, and we found that there were significant differences in body mass index, number of exercising days outside of cardiac rehab, HDL/triglycerides, waist circumference, and diastolic blood pressure in those patients who were in AACVPR certified programs compared to those who are not.

So that's kind of the overview of the rationale for these measures and the testing that we did to submit with these measures. I probably have a couple of colleagues on the phone, as well, from AACVPR. They probably should introduce themselves.

CO-CHAIR GIBBONS: All right. If there are additional representatives on the

phone from AACVPR, could they please introduce themselves?

DR. MARJORIE KING: They should be joining. There were two numbers that were circulated.

CO-CHAIR GIBBONS: Well, we are actually eight minutes ahead of time, so they may be not on the call yet and we may hear them come on. So we're going to move ahead with the first measure, 1496, safety standards. Leslie?

DR. CHO: So this is measure 1496. It's looking at cardiac rehabilitation program structure base measurement set of safety standards. And, basically, what it's talking about is that there needs to be a medical director present, emergency response team, minimal of BLS and ACLS training by one personnel, and a functional emergency resuscitation equipment.

I think the importance is self

evident. Patients are getting older. They're at higher risk. They have more comorbidities and medical supervision is crucial for good cardiac rehabilitation. About one arrest occurs in every 100,000 patient hours, so definitely safety nets are needed.

I mean, I think importance for that, and we can talk about scientific acceptability and whatnot later, but I think the importance is kind of self evident.

CO-CHAIR GIBBONS: Questions or comments about importance? Yes?

DR. JEWELL: So I think this is going to apply to all three -- all of the measures under this category. Dr. King made an important point that the information in the application, actually, under gap in care actually looks at gross rates of achievement in the certification program and a need for remediation before approval. It doesn't link to any of the specific measures that

are in here. So in other words, we don't know if the programs were denied or had trouble getting through the process based on these measures. That being said, I think the significance of all of these in terms of safety and efficacy of programs I think is probably the more salient importance feature, even though, typically, we're looking for a gap in care.

CO-CHAIR GIBBONS: All right. Other comments? I think we'll go ahead and vote on importance. All right. So 20 yes, one no. We'll move on to scientific acceptability.

DR. CHO: Right. So the scientific acceptability, I think all of us, I think, are aware that all these safety standard has to be in place for a good cardiac rehab program. My only concern, and maybe the representative can speak to this, is there's a growing trend for non-traditional cardiac rehabilitation, number

one; and that there's going to be also CMS funding for something called intensive cardiac rehab. And my greatest fear is is that those programs will have patients who are high risk enroll in them, and there will be no safety standards. And currently, because only 40 percent of the cardiac rehabilitation programs in the U.S. are certified, there's no real way to measure, as we've all alluded to. I think things like this have to be in place so that when home rehabilitation, non-traditional rehabilitation, and those intensive cardiac rehab/Dean Ornish kind of place -- things go into effect they still adopt a safety standard.

Just one comment. In your packet of all this scientific acceptability, you had listed for non-traditional CR that medical director will create a program for safety standards for those patients. Are we endorsing some kind of risk stratification

for certain patients going to home rehab versus hospital rehab in this document?

DR. MARJORIE KING: That addition of the phrase that the medical director would be responsible for setting safety standards for home programs was added in -- to answer your fears. I have similar fears in heart failure patients, for example, and the very sick and the elderly and the ones who have troubles getting in. I have very similar fears as a practicing cardiologist. And that was added so that there would be safety standards for home programs. One may say you do not do a home program for people who are at high risk, but you're not endorsing a specific risk stratification. You'd be endorsing the responsibility for setting up those policies and procedures. The responsibility is the medical director of the cardiac rehab program.

MS. LUI: This is Karen Lui. May I speak?

CO-CHAIR GIBBONS: Certainly.

MS. LUI: Okay, thank you. There's also, as far as the Medicare population, regulations in place for both ICR and standard cardiac rehab that require, as Marge said, the medical director requirements, the medical supervision requirements, the ACLS certification, and staff requirements as far as safety. So ICR is held to the same safety standards as standard cardiac rehab. And home setting is not supervised early outpatient setting currently.

CO-CHAIR GEORGE: Can the measure developers comment on how this data is selected, whether it's the physician in the facility that's saying, yes, I do have a procedure in place, or exactly how that data is collected.

DR. MARJORIE KING: The policy and procedures would be submitted online. The AACVPR certification process is an

online submission process to submit evidence that you are meeting the standards.

CO-CHAIR GEORGE: So the only way to attest that the standard is being met is through the certification organization?

DR. MARJORIE KING: Right. There is currently only one certification organization. It's not a money making phenomenon.

DR. CHO: You actually hit upon this major problem, which we'll come to in every single one of these measures, is that the only way we know is through the certification process for which only 40 percent of cardiac rehab programs are certified, so we have no idea. And this outcome study in Wisconsin between the certified program and the non-certified program, there is no hard outcome. There's no mortality outcome, MI outcome. It's all soft endpoints. It's a huge problem. I mean, I feel like we should be able to get

the data because CMS pays for cardiac rehab, regardless of whether you're certified or not.

DR. MARJORIE KING: Right. There is mortality data from CMS from the Brandeis group that there's a significant -- there's about 20 percent improved mortality in patients who attend and at least, I don't remember if it's 20 or 24 outpatient cardiac rehab --

DR. CHO: But that's not for certified versus non-certified, which is the main question here. You're measuring a group of patients. Your measurement is wholly dependent on AACVPR certification, but the question is is that certification -- can a program be just as good without being certified? Do you know what I mean? If you just don't want to do or you don't want to go through the paper certification process but you still follow all the guidelines, are you just as good?

DR. MARJORIE KING: And it's similar to the questions that were asked before. It's kind of a cart before the horse. If we don't have measures, then CMS won't test them, and then there's less likelihood for them to be tested, so it's something we're struggling with.

DR. RICH: I'm just thinking about this from a public reporting point of view, and I'm not exactly sure how we're going to report it. I know that in Michigan we tie our public reporting either to -- we do physician organization reporting or health system, but these can also be standalone. So I'm not sure, are we looking to do this just as a population measure? Because I think that's an important question, as well.

DR. MARJORIE KING: It's a per program measure, so per cardiac rehab program, whether they're standalone, based in a hospital or physician's office. So

it's --

CO-CHAIR GEORGE: Can NQF comment on whether you have other program measures in your other measure sets?

DR. WINKLER: We don't have a lot of them, but we certainly do have measures that are of similar nature. These are more structural measures. You know, do you have certain characteristics, yes or no, within your program. And while they're not large numbers, there are a few, so it's within the type of measure we have.

DR. RUSSO: And a similar question, too. Can you expand, I'm not familiar with the lit, that one study, how large a study was that comparing the certified programs versus non-certified, and then is there any way to even consider expanding the measure, to bring up the point to not just be at the certified centers so that we could measure that? Do we really know right now that the certified -- I mean,

it makes sense. It should be. And will it be required soon? Is CMS going to reimburse only the certified programs?

DR. MARJORIE KING: I will defer the question about CMS to Karen Lui, who is on the phone. It was 70 programs, 21,000 records analyzed for the Wisconsin data. There are plans for an AACVPR registry which will begin in 2012. But, again, that will be self-selecting for programs that want to participate in outcomes data collection.

DR. CHO: Maybe the CMS representative can answer this, but CMS pays for cardiac rehab, regardless of whether you are certified or not, and you have to have certain standards in place. Isn't that correct?

DR. MARJORIE KING: Karen actually is in AACVPR, but she knows the --

DR. CHO: But isn't that correct, that CMS will pay for that?

MS. LUI: At the current time,

there's no movement by CMS to reimburse only for certified programs. However, the organization is in the early planning stages of attempting to get dean status and moving that direction, but that's as far as it has gone to date.

DR. CHO: Right. So I guess what I'm asking is is there a way for us to get data from CMS to look at outcomes between people who are certified versus non-certified cardiac rehab programs to see if there's really a difference? You guys know who you certify, so everybody else is non-certified; isn't that true?

DR. MARJORIE KING: Yes. Does CMS do that for free?

DR. CHO: I mean, I guess the question is -- you know, I fully believe that you have a wonderful standard and I feel like everyone in America should adopt it. I am completely in agreement that these are excellent, and we're, at the Cleveland

Clinic, AACVPR certified, thank God. But the question is is that for these safety standards and for other standards, because you are only looking at people who are certified --

DR. MARJORIE KING: I know. We struggled with this. I don't know if Randy Thomas or Steve Lichtman are on the phone or can help me out, but we struggled and struggled and struggled and I hear you. And, unfortunately, you know, it went through my head we'll ask the Brandeis people, but the Medicare data they had for that was a decade ago.

MS. LUI: This is, this is --

DR. THOMAS: This is Randy Thomas. Can you hear me? This is Randy Thomas. I was just going to answer very quickly it would be ideal if we could work with the Brandeis investigators to answer that question, and we'd be very interested in doing that. Because we had a very short

time frame to put together the data that were presented to you, we weren't able to pull that off with the Brandeis investigators, but it certainly would be more accurate than looking just at certified program definitely.

And likewise when we presented to you the data from the registries from Wisconsin and Montana, those data are also somewhat limited because it's still a self-selecting group, even if they're not certified by AACVPR, but they've taken the initiative to be part of a registry, they get regular feedback, they work on quality improvement, et cetera.

So it would clearly be better to have from a Medicare general population of data to answer that question. We'd be more than happy to pursue that. We don't have the funds that would be necessary to do this on our own, but we're definitely interested in working with the Brandeis folks to see if

they would be interested.

DR. LICHTMAN: Yes, hi. This is Steve Lichtman. Can everybody hear me also? Okay. Yes, I think Drs. King and Thomas hit the nail right on the head. When you look at the evidence presented now, it presents a picture that you have to infer from. We have a published study from the Brandeis group we've all been talking about, and actually Dr. King understated the benefits of general standard cardiac rehab. It actually, depending on which regression model you use and which co-variate model you use, five years out after cardiac rehab, patients who attended cardiac rehab of the 28, a 54 percent decrease in mortality as compared to those patients who did not. Additionally, there's a clear dose effect with one visit being better than zero, 12 being better than one, 24 being better than 12. And what's kind of the standard full program of 36 being statistically better

than 24 in terms of morality/morbidity endpoint.

So we have that data to start with. Then we have the data for programmatic issues, certified and non-certified programs, why programs are certified, why programs are not. But, again, as Dr. Thomas said, that's a self-selected group. Certification is intended as a self-study process. We don't want to turn down folks for certification. We want people to go through the certification process to improve their performance measures so that, indeed, they do get certified and they have what we believe is a top quality program.

Then we looked at a database of 20,000 records, albeit from one or two states or a few states, that showed some differences between certified and non-certified programs. There were some differences with certified programs having

some outcomes that were better. Some outcomes everybody improved in. And, again, Dr. Thomas said that is a self-selected population, but even within that self-selected population there was some differential improvement for certified programs.

None of these hit the nail on the head. They're all inferential data where we can conclude cardiac rehab is good, certification is good, there's evidence that shows certification is better than non-certification. But, clearly, if we could identify certified programs, all of them, which we can, give Medicare those center identification so that we know where data is coming from and where bills are coming from, you can certainly construct a research study to look at Medicare endpoints, just as the Brandeis group did, mortality and morbidity data, looking at certified and non-certified programs, but that's going to take a while,

and we would have to get the buy-in from the research group at Brandeis also.

DR. JEWELL: So I think that that gets us in the right direction, that kind of approach. We're still challenged by the notion that even if there are significant differences between certified and non-certified, we may or may not know whether those differences are linked to the measures that are being submitted in front of us. And for me, that's one of the biggest challenges here. And, again, I want to disclose, as a member of that organization, this is an enormous struggle for me because I'd like to see them succeed. But when I put on my measurement hat, this is where I really am struggling.

CO-CHAIR GEORGE: Is there a cost to certification?

DR. MARJORIE KING: Yes, it's, roughly, $600.

CO-CHAIR GEORGE: Is that annual

or one time?

DR. MARJORIE KING: You then get re-certified every three years, so a couple of hundred dollars a year to a hospital.

DR. SMITH: So just a point of information, we are being asked to certify the certifiers, in a sense. So this process would result in ordaining the AACVPR with credibility. In order to do that, it would seem that we need the evidence mentioned to have some reason for doing it. I mean, I say I'm a strong supporter of rehab, underutilized, of major importance, but what are the criteria that will best allow us to say it is safe and where is the evidence that these criteria are valid? This process could be used to do that if it were modified in some way.

CO-CHAIR GIBBONS: Do the developers want to attempt to answer that point?

DR. MARJORIE KING: Well, we

don't own these measures. We wrote them. They're published, but anybody can use them once they're NQF endorsed. If someone wants to say, oh, I'm going to make money on endorsing cardiac rehab programs they can go out and do it. So it's not like --

DR. CHO: Right. But I guess the question, you know, I really want to say how much I'm in favor of cardiac rehab and all this stuff and how much I think the AACVPR does a wonderful, fantastic job. I think the thing that all of us are struggling with is because the measurement relies on certification, it is a completely circular process. Do you know what I mean? So what you really want to know is is, you know, somewhere in Alaska, non-certified CR program, are they doing -- do they have the safety in place? You will never know because they're non-certified. Do you know what I mean?

DR. THOMAS: This is Randy

Thomas. I was just going to answer very quickly. I agree with what you're saying. I think Dr. Lichtman had mentioned this that really the use of the certified data and the certified program data was really a surrogate to try to answer the question that you're talking about. And, clearly, the purpose of the measures are not to promote -- to promote the delivery of effective cardiac rehabilitation. And so the intent is to make this, these measures widely applied whether or not the programs are certified.

Unfortunately, with the relatively quick turnaround time that we had to put together data to show this committee, the best available data that we had was through the certification process. But we'll agree completely that that's really not the issue. Really, the issue is trying to implement these into programs, whether or not they're certified. So the tricky part

is try to identify the programs that are using the measures. The best surrogate we could find was certification.

DR. MARJORIE KING: Could I just say one thing? When we chose the measures back in 2005, I mean, I remember my heart sinking when the group said you have to write measures for cardiac rehab, thinking, oh, my gosh. And when we chose those measures, I was pleasantly surprised that the measures we came out with independently, independently came out with were very similar to the then existing AACVPR program certification measures, and it's because they're based on the core components of cardiac rehab, which are based on the ACC/AHA guidelines. You know, it's kind of mom/pop/apple pie that processes to do all these safety standards and these other things are kind of mom/pop/apple pie standards. So that part is circular, too, that the AACVPR certification process is

based on these measures because they're all based on common guidelines.

DR. LICHTMAN: Yes, hi. This is Steve Lichtman again. And just to echo what Marge just said in maybe a different light, I don't think we should also forget or disregard other measures of validity other than measurable quantitative data points. You know, these measures did go through an extreme process of content, context, and face validity analysis, and a lot of them, unless you did a very, very large scale nationwide study of the type we've been talking about, and even that might not show conclusive results, but a lot of these are really the result of committees, peer review processes, expert opinion, consensus, NQF-like scoring for each one of the points where measures were eliminated, measures were altered. So there was this multi-process, rigorous process of looking at the content, the context, and the face validity

of these measures. And as Dr. King said, you know, it was actually a marvel that these came about and that there was such good consensus on the final set.

CO-CHAIR GIBBONS: So my sense of the various comments made by the committee, if I can try to bring this to a halt and move to a vote, is that the concern is not about the measures but the linkage of the measures to certification by a particular group.

DR. CHO: Right. So if you have to be certified to -- in order -- if you have these safety standards and you're certified, you're going to be 100 percent. So if we go out and measure these people, it's going to be 100 percent because those people are the ones who got certified. So the goal of NQF is to identify people at gap, right? So you want to bring the rest into 100 percent. But if they're already 100 percent because they got certified

because there was four of these, then what have we accomplished?

CO-CHAIR GEORGE: And I'm struggling with how a rehab center that does not want to be certified for some reason, how do they report this data?

DR. CHO: CMS, you have to have certain minimal criteria in order to get reimbursement from CMS. So even if you're not certified, you have to have some kind of individualized planning, you have to have the code card and things like that. I mean, we should bring this to a vote because I think, you know, the issues are pretty evident.

CO-CHAIR GIBBONS: So we're going to vote first on scientific acceptability of the measures, okay? Scientific acceptability.

DR. JEWELL: Just to clarify, even though we've been talking about all of them, we are still just voting on --

CO-CHAIR GIBBONS: Safety. We're voting on safety standards.

DR. JEWELL: Right. But on this measure specifically.

CO-CHAIR GIBBONS: Right. This measure, 1496. David is getting ready. All right. We have 3 completely, 11 partially, 3 minimally, and 4 not at all. So now let us move to usability.

DR. CHO: Well, I think for usability you have to be certified in order -- it's currently in use for those programs that are currently certified. So if other programs want to join, then it would be in use. But then, again, they would be 100 percent.

DR. SNOW: Is there any way we could remove the need to be certified from this proposal and just look at how the various programs are doing with regard to safety measures?

DR. MARJORIE KING: You do not

need to be certified to use these measures. These measures were just -- the only way we had to test the measures was by testing the certification process. These measures could be used by anybody anywhere.

CO-CHAIR GIBBONS: They can be used, but where's the data which we have been requiring as part of the process for their use outside of certification? That's the dilemma we have.

DR. MARJORIE KING: Where's our competitor? Where's AACVPR's, quote/unquote, competitor.

CO-CHAIR GIBBONS: You're misunderstanding the point. Where's the data for the 60 percent of programs that are not certified. We don't have any.

DR. LICHTMAN: Right. But that's almost the point. The point is that programs choose to be in certification. If performance measures that set the standard for cardiac rehab were passed, and the goal

of AACVPR and everybody in cardiac rehab is to have quality programs across the country, it would at least -- it would provide standards for programs that choose not to be certified to be able to achieve to set their standards to, so it would actually increase the number of quality programs, we believe, across the country. We're not looking at certification as part of these performance measures. These performance measures stand alone, and they hopefully will provide the standards that cardiac rehab centers have to achieve.

MS. PACE: Just one comment. We do have measures that have been endorsed that come in initially from a particular program. Registry measures are an example. The goal is that by the time of maintenance review they're publically reported and more widely used. So that, in itself, shouldn't prevent it from going forward. But then the question would be are these appropriate

performance measures? Structural measures have other issues.

CO-CHAIR GIBBONS: At a minimum, presumably, time limited endorsement to demonstrate collection of data outside of the program.

MS. PACE: Our testing doesn't require widespread national testing. Even other measures do it on a sample, and so that wouldn't necessarily limit it. I think the broader issues are the structural measures.

CO-CHAIR GIBBONS: I think we have to have a vote on usability. The vote is 2 completely, 12 partially, 4 minimally, and 3 not at all. And now feasibility.

DR. CHO: So I think it's very feasible if they're certified. It's not that feasible if they're not certified, unfortunately.

DR. MARJORIE KING: But there's nothing that precludes, I'm sorry for

butting in, but there's nothing that precludes a non-AACVPR member from applying for certification.

DR. CHO: I'm so with you on the certification. I feel that, you know, the thing is is I have existential angst about this thing because I totally believe in the certification process. But the point is not the certification. Do you know what I mean?

DR. MARJORIE KING: I know, I know. And we struggled with this, Randy and I and Steve and Karen struggled with it.

DR. JEWELL: So you made a comment earlier that I just want to revisit. You will be the measure -- if we were to endorse this you would be the measure stewards, as opposed to anybody else using them. That's a different issue. You would be the measure stewards. So the testing that they're talking about and the ability to gain data, whether it's from CMS or wherever, on the currently non-certified

programs falls to AACVPR to do.

DR. THOMAS: Excuse me. This is Randy Thomas. Actually, this is a measure that's jointly developed by the American College of Cardiology, American Heart Association, and AACVPR. Just like with some other measures we developed together, we'd be joint stewards.

DR. JEWELL: Okay. And thank you for that clarification. I think my point was simply to make sure. You've talked a lot about the challenges you face with testing on the front end of this. I want to make clear that everybody -- that you all understand that if were to conditionally or in some way endorse this and require testing that that means that it's up to you to go to CMS or wherever else you think you can get the data in order to come back and demonstrate that this measure has utility. So we hadn't touched on that yet, so I just wanted to be clear that you understood that

that's the implication of a decision like that.

DR. MARJORIE KING: We know.

DR. AYALA: I have a question, too, for those that are not certified. With the first point of the numerator statement, is there a set of standards that, like I'm just trying to understand how a reviewer or how the institution would prove that they actually have appropriate policies and procedures. You know, it's a very broad term. I was just wondering how specific is that in terms of the content of those policies and procedures, and how would the institution report that and how would it be monitored and audited?

DR. MARJORIE KING: So you're saying if it's outside of the AACVPR certification process which has trained reviewers with templates and best practices and all that sort of thing. Somebody else help me out within the group. I mean, it

would be like any process when you're looking at data.

DR. THOMAS: This is Randy Thomas again. I would think it would depend on the application of the measures. For example, if a health care system implements these or a third-party payer institutes these measures, they would be the ones who would be collecting and auditing the information. If Medicare, for example, were to adopt these as part of their evaluation of reimbursement for cardiac rehabilitation, we would assume that there would be a process in place to help audit and identify the adherence to the measures. The certification process through AACVPR is one mechanism to make that happen, but it would really depend on the organization that is using the measures.

DR. MARJORIE KING: And as I look at this numerator statement, it's fairly explicit about policies and procedures

consistent with evidence-based guidelines, safety standards, regulatory standards. Those are in the literature.

DR. AYALA: Right. I understand that. I'm just thinking of the feasibility of reporting on the details of that and on the auditing process.

DR. MARJORIE KING: Right. It would require similar to the AACVPR process with training and --

MS. SZUMANSKI: I think there is one piece in here that is missing from the safety standpoint, and that is the concept of hand-off communication which has been shown nationally to be a major problem.

DR. CHO: It comes up at different --

DR. MARJORIE KING: We have a specific measure for that.

CO-CHAIR GIBBONS: Okay. I think we need to vote on feasibility. Okay. The vote is two completely, seven partially,

eight minimally, and three not at all. Now, the final vote: does the measure meet all of the NQF criteria for endorsement? The vote is six yes, fifteen no.

So we'll now move on to 1494, cardiac rehab response to therapy. And I presume that the construct here is very similar. Ann, why don't you guide us through this one?

MS. DE VELASCO: I sort of feel like I'm trying to guide you through the mud because I'm kind of befuddled and muddled by all of this anyway. But this program actually, this measure is related to the monitoring response to therapy and documenting the program effectiveness. It has to do with written procedures and policies, and it is a patient-centered and structure-management type of program.

The importance, I think, was alluded to in the comments prior to this. Obviously, the effect of cardiac rehab on

the patients' survival and quality of life is well documented. And the only thing I saw was that there was some performance gap when they referred to the people who were AACVPR certified that it wasn't, you know, the proportions were disproportionate. The other thing that I saw that was, the percentage of patients to referral, it said that 55 percent are referred but only 19 percent actually enroll. So there was a performance gap, at least at that level. However, the importance of the measure seems to be solid. There are some disparities, though, that it is less prescribed for the elderly, the women, and minorities. And, obviously, in order to implement all the procedures, you have to have some type of recording process and documentation, which this measure specifically addresses.

CO-CHAIR GIBBONS: Are there additional comments before we vote on importance? Okay. Let's go ahead and vote.

Twenty yes, one no. Okay. So now we'll move on to scientific acceptability.

MS. DE VELASCO: Okay. As far as scientific acceptability, it has been precisely specified. The numerator has four components that are listed on the measure to document the percentage of patients who have received a formal request to cardiac rehab; also document the standard plan to access completion of the prescribed course of cardiac rehab; document for the patients a standard plan to access certain outcome measures at the initiation and at the end of the completion of cardiac rehab, and the outcome measures are actually outlined in the AACVPR performance measures; and also to describe the programs' methodology to document program effectiveness and initiate quality improvement strategies. This is per reporting year, and the denominator is all cardiac rehab programs, male and females, 18 years or older. Again, the time element is

per the performing year. There were no exclusions and no variables. There appeared to be no risk adjustment, and the data source can be paper medical records or computer provided. It also has relations to the organizational plans and policies that are implemented by the cardiac rehab program using their departmental records. For the people that are AACVPR certified, there is an outcomes data registry that is going to be collecting and analyzing the data, as well.

So that, basically, reliability, the reliability testing has been done through the AACVPR, which, again, is this kind of circle of things that we did before. And AACVPR is an all-or-none phenomenon. You are not partially certified. And they have an extensive review program in place so that they can remediate people who are not approved and can come back for certification.

Validity testing was done through peer review and extensive record review. And validity testing was also completed without any obvious outliers that I could tell.

So in February 2011, they were going to make additional testing records available to us. I don't know. I didn't see those records. But there were no exclusions in this section.

Again, the difficulties I had with the scientific acceptability was the circle within the AACVPR certification, if we were just basically certifying a certifier. But that's been addressed before, so I don't have any new insight on that.

CO-CHAIR GEORGE: When I looked at this, there's four components in the numerator and three of those are really patient level, one is a system level. And I'm wondering if it wouldn't be possible to

go back with the patient level components and construct measures that really address each of those items on patient level.

DR. MARJORIE KING: Are you referring to does the patient get in program, finish program, and improve their outcomes? And then the reason we wrote the fourth component was, okay, so you document all that, but your patients don't get in program, they don't finish program, and they don't improve their outcomes, then what do you need to do about it, have a performance improvement project to improve those. So that was the thought process behind the measure.

CO-CHAIR GEORGE: I think, from a measurement standard, the first three could be reconfigured so that you could report on those components individually that go into each one of those and come out with something that is meaningful, perhaps at the patient level, for the rehab unit.

DR. MARJORIE KING: And these measures are actually being used in non-AACVPR places, like these registries which are not part of the national registry. The Wisconsin registry is not part of the national registry, and Montana also uses these measures. So these are not, these are used outside of AACVPR.

MS. DE VELASCO: One thing that I thought was interesting is that risk factors that influence outcomes should not be, obviously, exclusions, but that patient preference is not a clinical exception to eligibility because it could be influenced by provider intervention. And on a day-to-day basis we certainly see whether patients come to cardiac rehab. There's a huge provider intervention component to that of how attractive you make it to them when you basically sell the program on the phone and how diligently you follow the patient from admission through discharge and then at home

to get them back into the system. So even though that's not an exclusion, it certainly does play a part in how many people actually do attend cardiac rehab.

DR. MARJORIE KING: And that's another part of why this measure was written is to hold cardiac rehab programs accountable for working with the referring physicians to get the patients from those endorsed referral to cardiac rehab measures actually into and finishing program.

MS. DE VELASCO: Right. Which speaks to the hand-off that Kathleen mentioned is interdisciplinary connections with other providers and different disciplines that the patient would be subjected to after their diagnosis.

DR. MARJORIE KING: Right. Our referral measures do have communication embedded in them.

MS. DE VELASCO: Right. And I think that's the key to a successful rehab

program, too, is to have the interdisciplinary approach.

MS. SZUMANSKI: Your reviewers that you describe in your analytical method, are those individuals permitted to review their own program?

DR. MARJORIE KING: Absolutely not.

MS. SZUMANSKI: And do you send more than one reviewer for review, similar to what AATB would do for bone marrow?

DR. MARJORIE KING: It's a paper-based review currently with plans to do on-site reviews. And if there's any controversy, the chair of the committee does the review as well.

MS. SZUMANSKI: So this is strictly a paper review and not an onsite --

DR. MARJORIE KING: Or an electronic review. Yes, it's not an onsite review.

DR. AYALA: Would it be

appropriate to have an exclusion for the first component of the numerator if a referral request was received but that the medical director who reviewed the case decided it was inappropriate to have the patient enroll in the cardiac rehab program?

DR. MARJORIE KING: Well, there are medical contraindications to referral, things like very severe aortic stenosis and very severe hypertrophic cardiomyopathy, but they're very rare. And severe dementia, obviously, would not be appropriate with lack of carryover. But there's very few patients who don't benefit from cardiac rehab.

MS. DE VELASCO: But I think, in reality, there are certainly cases where patients do not get accepted to cardiac rehab because they don't fit within the construct of an outpatient cardiac rehab program. They may be referred to cardiac rehab and may be totally immobile. They may

be in a wheelchair or there may be other things, so they would absolutely --

DR. MARJORIE KING: I work in an inpatient rehab hospital, so, to me, no one is not rehabbable. I'm sorry. So I have a different view. There are adaptive modalities and things you can do with people, so I may not be typical. Randy may want to address that.

DR. LICHTMAN: Physical disabilities would never preclude somebody from participating in cardiac rehab. In our facility, we even had an individual with quadriplegia who participated using the shoulder muscle, so I think without the extreme medical or cognitive exclusions, I've been doing this 20 years and I've never turned down a patient for physical disability. We work around it. Some may not progress as well as others, but we certainly would enroll them.

MS. DE VELASCO: That's good to

know. I'm glad to know that. Our experience is a little bit different, but we have something to look forward to then to achieve.

DR. MARJORIE KING: Right. We want to drive enrollment and adaptation of programs and that sort of thing.

DR. KOTTKE: Yes, the old Kottke back of the spreadsheet analysis here. The impact of getting everybody into cardiac rehab would be about four times the impact of angioplasty and everybody immediately for STEMIs. We talk about setting clocks on ECG machines and all those kind of things, and I think where my angst is about this is that CMS collects all the data on, for example, aspirin, you know. And if we turn this down, you know, we're asking them to find somebody who will collect all the data.

It's like this problem of why don't we talk about diet? Because we don't know how to talk about diet. Nobody denies

that nutrition is very important. And I don't think we have the answer, but I think we need to, somebody needs to really work on this answer. And I don't know if it's CMS to adopt stricter criteria for payment or something, but I'm very uncomfortable here because the impact is very large. We risk shutting down a very important process that patients clearly benefit from, but I haven't figured out the solution.

CO-CHAIR GIBBONS: All right. I think we need to go ahead and vote on scientific acceptability. Okay. Completely three, partially fifteen, minimally three. Usability?

MS. DE VELASCO: We didn't have any problem with the usability. It is currently in use, and it's publically available on several websites. Let's see. We have recognized the expected outcome of these cardiac rehab programs. The reporting is done, and it's harmonized with the other

measures that we're reviewing for cardiac rehab. We think that it encourages cardiac rehab secondary prevention programs to collect and respond to these outcome data and that to improve enrollment and completion of cardiac rehab. It also stimulates performance improvement strategies for cardiac rehab professionals - sorry, yes, if you are certified. So in agreement to what Dr. Kottke says, of course we heartily endorse cardiac rehab and the benefits to the patient. It's getting to that point that is what we're trying to debate here today, I think. And as far as usability, it appears to be something that definitely meets the criteria of usability.

CO-CHAIR GIBBONS: Any other comments before we vote on usability? Okay. Let's go ahead and vote. Seven completely, eight partially, six minimally, reflecting the difficulty of this dilemma for sure. Moving on to feasibility.

MS. DE VELASCO: Okay. As far as feasibility goes, the only thing is that if the patient fails to complete the program it may affect the program's ability to capture the individual outcomes and accurately reflect the program effectiveness. But we're aware of the fact that attrition is a challenge in cardiac rehab programs where self motivation is a significant problem. However, the feasibility could be affected by closely monitoring the barriers to completion and not waiting until the very last minute for when you think the patient is going to, for example, complete 36 visits if you try to proactively collect some of the data prior to the end of especially the patient component data that they have to respond to.

They conducted this work group in the Wisconsin study, and the refinements were made to all the different completion reasons or the reasons for not completing

cardiac rehab. Those things are being tracked.

So far as the feasibility, it seems to be a lot of that is already in place. The cost seems to be minimal. We reviewed the costs that belong to AACVPR, which currently is part of a lot of this stuff in this measure. So for a relatively low-cost process, we can have a significant impact on the outcome in patients who require cardiac rehab as far as their quality of life and morbidity and mortality. So we think that it's feasible as it is right now, although it's dependent on the AACVPR and how that works out.

CO-CHAIR GIBBONS: Okay. I think we'll go ahead -- oh, sorry.

DR. JEWELL: Just to clarify, this measure is asking whether or not a program has a policy to do all these things, yes? It's just simply a yes/no question.

MS. DE VELASCO: Yes.

DR. JEWELL: Okay. Because the way you're describing it it sounded like more than that, and I was thinking, God, I really missed something when I re-read it.

MS. DE VELASCO: No. I was just adding, that's sort of the endorsement of cardiac rehab stuff I was adding, but it basically just refers to the policies and procedures.

DR. JEWELL: Okay. So just to be clear, it's a measure that says does the program have a policy or not, and it's verified through certification, as was the prior measure. Okay.

CO-CHAIR GIBBONS: Tom?

DR. KOTTKE: I'm just trying to figure out the logic here of this. Let's go back to 30-day mortality post-MI discharge. CMS only has data for Medicare, and we didn't put the kibosh on their indicator because they couldn't tell us anything about people under 65. And so if this indicator

or this measure is for cardiac rehab programs that participate in AACVPR certification then it seems to me that it's okay. I mean, yes, we don't know about the other half of the glass, but other measures, too, we don't know the entire universe of patients with the condition.

DR. JEWELL: Yes, but just to be clear, we actually chose not to endorse or re-endorse measures because there was no demonstrated gap in performance. Based on your math that was 98 to 99 or 100 percent wouldn't save many lives or do many things for other measures. We have no idea if there's a gap in performance here because we don't have beyond the remediation efforts of the overall certification whether this particular issue is a problem.

DR. KOTTKE: But we have no idea whether there's a gap in performance between Medicare age and non-Medicare age and 30-day mortality.

CO-CHAIR GIBBONS: Dana?

DR. KING: Tom, point very well taken. I see a bit of a difference between those two things. One is if we were talking about cardiac rehab and we said, well, we only have information in 33 states out of the 50 and we're working towards getting all 50, or if we said we only have information on people over 50 years of age but not under 50 because of some unusual thing. But it would be a difference if we said we only have information in PCI mortality on people that didn't die but not on the ones that did. And so that is not the same thing. We can't say we only get information on the ones that are doing good and not on the any of the ones that are doing bad. I mean, all of us are like throwing ourselves onto knives, but right now it's bad data in and bad data out. I mean, the measure is good, but one of the requirements of NQF, and everyone is up there shaking their head, is

that you have to have some data to show the utility of the measure and we don't have the utility to show the good and the bad and we don't have a good picture. So that's the difference. It's not an arbitrary or a sampling problem. You have to meet the measure before you enter data, and it's not the same.

CO-CHAIR GIBBONS: I think we should move on to a vote on feasibility. So we have one completely, twelve partially, four minimally, and four not at all. So final vote for this measure, 1494, response to therapy, does it meet criteria for endorsement? Okay. We have three yeses and seventeen nos.

So we're going to move on to 1497, which is cardiac rehab risk for adverse events. And, Dianne, you're the primary reviewer.

DR. JEWELL: I am. So this is a measure that looks to assess the presence of

two assessments of risk for adverse cardiovascular events in newly-enrolled patients in cardiac rehabilitation. All the previous comments, at least that I've made, stand here. It's essential to know not only because we don't want our patients blowing up on our treadmills but also because we want to optimize the effectiveness of the care they get, and part of that optimization depends on understanding their risk for events.

So absent data that links a gap or demonstrates a gap on this measure, I would urge us to vote yes on the importance because I think, at a minimum, that needs to be stated somewhere in the record that we think that this is, in fact, important for all programs to do.

CO-CHAIR GIBBONS: I would suggest that much of the discussion for the previous two measures does apply here, as well. So any other comments about

importance before we vote on importance? Let's go ahead and vote. Nineteen yes, two no. Scientific acceptability.

DR. JEWELL: I am torn because, you know, the cardiac specialist in me definitely wants to see a higher standard of performance in the programs and wants to be able to reflect that in the quality measurement efforts. The measurement person in me says that all the things that we've talked about make this measure not ready for prime time across all three domains. That's just the frustrating part. That's where I am.

CO-CHAIR GIBBONS: Anybody else want to share any additional existential angst or otherwise? Leslie, it's a term I'm now going to carry with me and I'm going to associate with you. It is a perfect succinct summary of the problem.

CO-CHAIR GEORGE: I would just say that this measure, again, allows itself

to be easily reworked into something that I think would be much more usable and would encourage the measure developers to continue to work on that.

DR. AYALA: I'm going to ask the question that if we had data that showed mortality and stratified it by whether or not an institution were certified or not and if that showed a difference, a significant difference, would we even need these measures if we know that those that got certified would meet all these measures? So if we had that one measure of mortality stratified by, you know, whether or not the institution was certified, then wouldn't that be enough for CMS, for example, to say you've got to be certified and then these measures wouldn't even be relevant because they would all be met by the certification process?

DR. RICH: Maybe you would want that measurement, that certification in

conjunction with the composite score so the patients would have some idea of differentiation between places.

MS. SZUMANSKI: One could use that same argument related to the Joint Commission that if you are accredited you must be meeting all of the elements of performance that they define and, yet, we know nationally that is not the case. So while it sounds good, it ain't so.

DR. RICH: That's why throwing in a composite score would give people a little bit more information.

CO-CHAIR GIBBONS: I think Kathleen's point is an outstanding one. It gets to the heart of what we all know goes on in terms of certification processes.

DR. JEWELL: Sure. Although that being said, at least to the extent that these measures are structural so the presence or absence of the policies, as opposed to the actual effect of them. And I

agree with you, having lived through the Joint Commission myself a few times, it probably would be more true perhaps but still with its limits.

CO-CHAIR GIBBONS: Okay. I think we want to vote on scientific acceptability. Okay. One completely, thirteen partially, six minimally, one not at all. Moving on to usability. Any additional comments on usability?

DR. JEWELL: Sorry, no. I've meant my comments to address all three.

CO-CHAIR GIBBONS: Okay.

DR. JEWELL: Because the same issues apply.

CO-CHAIR GIBBONS: We'll go ahead and vote on this one. Okay. So two completely, ten partially, seven minimally, and one not at all. And feasibility. I propose we just go ahead and vote. Okay. So we have eleven partially, eight minimally, one not at all. And now can we

move to the final vote on endorsement? We have two yes and nineteen no.

All right. So now we're going to move to 906, the cardiac rehab composite. And Leslie is once again the reviewer. And, Leslie, I guess my first question to you is whether we should follow David's previous motion here with respect to the voting on this measure since it's a composite of the three we just talked about?

DR. JEWELL: I think we should follow David's example.

DR. CHO: Can I just make one comment, and that is I think all of us in this room are truly angst-ridden because of the fact that we're voting this down. But I just want to commend the AACVPR for all their hard work. I think if somehow these were all aside from the certification process, we would wholeheartedly endorse them 21 to zero. And I think that just because we voted this down does not, in any

way, shape, or form, reflect the fact that there has to be standard in America for cardiac rehab programs.

CO-CHAIR GIBBONS: Well stated. The Chair would suggest that we take the -- the votes were slightly different, so the Chair would suggest that, for the record, we take the last vote on 1497 and enter it on the composite, if there is no objection.

DR. AYALA: I just had a question on the math on the composite. Is this pretty straightforward, or does this got some weird math in it?

DR. CHO: Unfortunately, it's got a little bit of weird math, but it doesn't have the composite like in the AMI. So what it is is that you have to fulfill 11 of these or 10 of these composites. And it's purely based on the certification process. They have to come up with an assessment. Depending on some of the composite, the compliance is anywhere from 60 to 90

percent, like, for instance, LDL and all this other stuff. I think that there are some questions with composite only because the blood pressure definition has changed, as Dr. Smith had alluded to, and also the waist circumference is purely Caucasian based and not ethnically based like for Asians and whatnot. But I think that the problem with this is so similar to the other ones that I think the voting should be the same.

CO-CHAIR GIBBONS: Okay. So hearing no objection, we'll enter the vote for 1497 for this measure. Now, my sense of this is that everybody is as uncomfortable as I am at the fact that we have voted these four measures down, so I think we need to record something for the record for the measure developers and for others that we're not against cardiac rehab. And I would encourage others to weigh in, but the issues I see are, one, linkage to certification

and, two, absence of data for centers who are not certified with respect to performance gap and validity of the data, that those two are the fundamental flaws and that if they could be addressed by the measure developer the likelihood of favorable review would be greatly enhanced. They are structural measures, but I think Leslie has nicely outlined the need for quality improvement in cardiac rehab in the United States.

Are there other fundamental flaws that we should mention?

DR. KOTTKE: My feeling is that they could, it would take some time and I think there would be foundation money available to look at CMS data for outcomes by certification, and that would satisfy the panel that there's a problem and that the attributes of certification are important to promote.

CO-CHAIR GEORGE: I would also

stress that I think that it's important that we have some patient-level measures in this set for cardiac rehab, in addition to structural measures, to consider.

DR. SNOW: Question. How long do you think that process might take? Well, yes, just a guess, you know, getting foundation --

DR. KOTTKE: I would have to ask somebody who has actually worked with CMS data. I don't know.

DR. SNOW: Well, okay. That kind of answers the question. That could be a slow process because it probably won't come back to this committee is what was in my thoughts, even though --

CO-CHAIR GIBBONS: It would be highly unlikely to surface in this year's --

DR. SMITH: So the proposers should be encouraged to be more inclusive of programs in this project and not link them specifically to one form of certification,

right? That's what we're saying.

CO-CHAIR GIBBONS: Helen?

DR. BURSTIN: Just one point. In the spring we actually did endorse two cardiac rehab measures, our first ones, which was patient referral from an inpatient setting for cardiac rehab and patient referral from an outpatient setting for cardiac rehab, both of which were endorsed at the patient level. So we do have two relatively new ones, so I think there's something to build on. But, again, agreed. Particularly, patient-level measures, I think, as Mary has pointed out, would be ideal. We just had discussion about sort of rules of the road for measure construction, and yes/no structural measures about specificity are just not going to make it through any panel kind of going forward. So it's good to start thinking about how to make those measures more robust as long as you have the time.

CO-CHAIR GIBBONS: Okay. So I think we want to, I'm trying to get a crisp sort of message back to the developers that this committee is comfortable with, and I think we've outlined four components: the absence of data on patients who are not certified or the linkage of these measures to certification, the absence on data of patients who are not certified. Tom has highlighted the absence of outcomes, favorable outcomes related to certification, and then, finally, the need for patient-level measures.

I think that's a pretty comprehensive message with those four components to indicate that we believe that if they can be addressed that rehab measures are important. Can I have a show of hands in favor of that? Okay. Hands down and any nos? Okay. Is there anything else that anybody can think of to satisfy their existential angst before we take a brief

break at this point?

DR. WINKLER: Public comment first.

CO-CHAIR GIBBONS: Public comments. I'm sorry. Public comment? Okay. We will take a 15-minute break. And those who need to leave, thank you for your diligent service during these two days.

(Whereupon, the foregoing matter went off the record at 1:23 p.m. and resumed at 1:45 p.m.)

CO-CHAIR GIBBONS: So --

DR. SMITH: Ray, Ray?

CO-CHAIR GIBBONS: Yes.

DR. SMITH: I have some information, if I could.

CO-CHAIR GIBBONS: Absolutely.

DR. SMITH: It turns out that existential angst is defined, and I'll just, within a minute, I'll let you know. It's one of the three corners of the existential triangle, along with People as Scenery

theory and the Anthropic Fallacy. People suffering from existential angst are either not convinced that they exist, unsure why they exist, or not at all convinced that anything really exists. It goes on with other information, but I thought that would reassure you.

CO-CHAIR GIBBONS: I think that's terribly reassuring, Sid. I feel so much better already. I think we ought to send Leslie an e-mail. Staff, you're hereby directed to send Leslie an e-mail that this was defined in her absence.

DR. SMITH: I talked to her before she left. I confronted her. In fact, that may be why she left. I'm not sure.

DR. MAGID: You know, I was also thinking that since Bruce is done with his training maybe we could go back and say that measure is not necessary.

CO-CHAIR GIBBONS: All right. So

we are now faced with a series of difficult issues. We will not get through them today. We will just get started on them today. But they deal with sort of insights that I think we've now gained from this process, from the difficult issue of competing measures that we've identified in several constructs, and from the issue of harmonization.

So several of you have spoken to me about this one, and I must admit, as I read through the applications, it truly bothered me, and that is just healthcare disparities. It was my feeling, as I looked at the applications, that a number of applications sort of chose to disregard that field. One or two has listed it as not applicable, which I found unbelievable. Others just put minimal, if any, answer in that field. And in response, several of our developers stated something to the effect of they had that data but had never looked at it.

I think we all work in environments that pride themselves on trying to deliver the same care to every American, regardless of race or gender, and this notion that disparities are not important when they've been documented in our healthcare system for at least 30 years and been the subjects of two IOM reports I just find totally unacceptable.

So what I would propose are two things. Number one, in the short term, we can ask the staff to convey and measure back to the developers that we actually expect to see that data that they said they have and they're going to get because I think now is the time to sort of push them to actually look at the data with respect to this issue and we make it clear that we expect to see that back because I just don't think we can let it fall through the crack.

So that's the short-term first step we can take and invite discussion or

comments about other short term things we can do in the next few months before I address the longer term issues. Are there other things that have occurred to people that they are concerned about?

DR. RUSSO: Just to clarify that one, are you implying that the initial application should include any literature that's currently available regarding disparities, plus when --

CO-CHAIR GIBBONS: That's the long term because we can't change the initial application for this round. It's already happened, so that's the long term. But in the short term, for the people that said, oh, we have that data but we haven't analyzed it, which I find just unacceptable, totally unacceptable for somebody who works for an agency of the United States Government. And I don't think they would want that in the public domain on the front page of the Washington Post. We have the

data but we haven't looked at it after three years? Come on. I mean, I know there are limited resources, but this is so obvious. And that's one of the reasons that disparities persist is people don't shine the light on them.

So any other ideas about what feedback we can give in the next few months?

DR. THOMAS: You know, another thing I noticed was some of them just said no disparities, so there were those that said, you know, we have it and, you know, we just haven't addressed it and there was others that said no, so just to make sure that those organizations, as well, address that issue because that's unlikely, you know.

CO-CHAIR GIBBONS: Right.

DR. AYALA: And in addition to the race, ethnicity, and gender, I would add payer source because I see that they actually have that data, as well. And age,

someone mentioned age, as well.

CO-CHAIR GIBBONS: Other thoughts?

DR. MAGID: Yes. I would just encourage the measures that, and we talked about this, the measures that talk about prescription of a medication at discharge, to really think about the possibility of going beyond that, because we know that a fair number of patients, you know, they may write that the patient is going to get this at discharge but then we find out later they either don't pick it up even the first time or refill it. There's a huge drop off at the second. So being able to look at these medications, which are really intended for chronic use over a time, make sure that they actually receive it and that they take it for longer. I know it may not be something that we can request them to do right now, but the next time they come around that's probably where they need to go.

CO-CHAIR GIBBONS: Okay. So I'd like, I think there will be more authority in this if there's a vote. So in terms of the short-term communication regarding the need for data regarding disparities, can I have a show of hands as to all of those who are in favor? Okay, all right. And in terms of David's suggestion that we encourage measure developers to look at medication issues over time rather than a single point in time, those who are in favor of that? Are there any nays?

You know, unless you prod them they're never going to try. All right. Now I want to move to the --

DR. BURSTIN: A quick follow-up question as long as you're being so decisive. So on the disparities piece, so, yes, you want to see that they include data on known disparities. The second thing is if you guys come forward and say this is a measure of known disparities, part of what

we're struggling with is should the measure we currently tag it as disparity-sensitive and indicate it's a measure that should be stratified, would you like to see that stronger --

CO-CHAIR GIBBONS: Okay. That's the longer term so that's about the next part of this.

DR. MAGID: Actually, I think the interesting thing about disparities would be to look at to what degree the disparity occurs because patients of certain characteristics go to hospitals that perform worse versus actually within a hospital whether they receive it, so sort of Betsy Bradley's work where she showed that the disparity between, say, African-Americans and whites was more due to the hospitals that African-Americans got cared for than to actual disparities within a given hospital.

CO-CHAIR GIBBONS: Yes?

DR. RUSSO: Just another comment,

since you're asking for other, is there someway that there could be some mention of what expectations would be if the performance, as they're resubmitting, if the performance on the measure was greater than, it's being considered for retirement so it's greater than a certain percentage and there's no variation across anymore, is there some way we could suggest before it's submitted, some standard that maybe --

CO-CHAIR GIBBONS: Well, there's another group already tasked with that. So I think that already exists. Dr. Masoudi?

DR. MASOUDI: Yes, I was just going to say I completely agree with what you're saying, and I think the measures that I discussed actually include it, and that's not the point. The point is, from a developer's perspective, it would be good to know in the submission form exactly what is being expected. People are interested in age, people are interested in gender, if

they're interested in race, whatever the stratification variables are, it would be good for developers to know that that is being mandatory as part of the submission and what the stratification variables are.

CO-CHAIR GIBBONS: I wholeheartedly agree, and that was going to be my longer term suggestion that the blank, and help me what field it is, 2G, H? 2H. That that blank is going to be given high priority and that the submitter needs to supply data regarding race, gender, age, payer. What else did we mention?

DR. AYALA: You have to include ethnicity because the OMB --

CO-CHAIR GIBBONS: Ethnicity. And if they can't supply that data, indicate why. But that's a requirement. I would suggest we go on record empowering the staff for the future applications that unless that blank is adequately detailed we will not consider the application. That's the only

way I think we can drive the process. And as I said, I think this is an important societal issue that somehow or other is getting lost in the field of measurement, and I think that's appalling.

DR. SANZ: I wish to add rural/urban since, according to our Chair, 25 percent of the people in the United States are in a rural area. And these measures, many of them have significant differences in rural areas.

CO-CHAIR GIBBONS: Now, sometimes all of these things won't be available, but we'll at least hold people's feet to the fire to try to get them. Dr. Masoudi, do you have other suggestions?

DR. MASOUDI: Just one other quick thing, which is that all the measures for this cycle have already been submitted.

CO-CHAIR GIBBONS: Oh, yes.

DR. MASOUDI: So they won't --

CO-CHAIR GIBBONS: No, they're

not affected. This is for the future. This is for the future.

DR. MARJORIE KING: Easily go back to those who have --

CO-CHAIR GIBBONS: Oh.

DR. WINKLER: Measure developers can expect that measures that will be viewed in phase two we'll be back at you.

DR. AYALA: I mean, in answer to Helen's question, there were three specific indicators that we talked about here where people raised the question about race and gender, and they were 0289 median to ECG, 0163 primary PCI within 90 minutes, and 0290 median time to transfer to another facility.

CO-CHAIR GIBBONS: I think it's wonderful that you kept track of that. I really do. I commend you for keeping track of that. So now we get into this question that Helen raised, which is not actually about blank 2H. It's about the earlier blank.

DR. BURSTIN: It's still about disparity. So the question would be we currently can tag them with the three that you just read off. We'll say these are disparity sensitive, and we would encourage these measures to always be stratified. Should it be stronger? Should it be that these measures should not be publically -- I'm just throwing this out, just curious here, you know, should measures like that where there are known disparities always be stratified so that you don't ever just look at a lump without being able to see what's underneath it?

CO-CHAIR GIBBONS: Thoughts about that issue?

MS. ALLRED: I would say absolutely yes because we already know from our experience that there are a lot of disparities that exist, and it isn't just women versus men. Sometimes it's young women versus older women. Young African-

American females, we know the first thing that is looked at in them when they go to the emergency room is cocaine use and instead of, you know, chest pain being heart attack. So I think there are a lot of issues that could be looked at, and that information ought to be readily available on any of these that are coming off the medical records.

DR. KOPLAN: I think everything that's been said is extremely important, but we just want to be a little bit careful not to be viewed as, I would think, as a group where that's like our only first priority because there's some other aspects to measure development.

DR. BURSTIN: There's a whole disparities group that's beginning in a couple of months. I was just really, since Ray brought it up, picking your brain because that's exactly what they're going to have to figure out.

DR. KOPLAN: Yes. But to make it a requirement across the board is a little -

CO-CHAIR GIBBONS: Well, the requirement is that they not blow the blank off, which many of the submissions clearly did. They put nothing in there and, yet, when questioned here said, well, we have the data, we haven't analyzed it. I'm sorry. They should have analyzed it and put it in the blank. So that's the driver, Bruce. And sometimes they didn't have the data. It's not a field that they're capturing but, hopefully, it will inspire them in the future to add that in the next round or in the next update or whatever because, you know, we've struggled for too long as a country without very much progress, and I think we need to start to make some progress.

DR. AYALA: You know, one positive thing about linking quality measures to disparities is that it's been

shown, at least in the Robert Wood Johnson Foundation's Expecting Success Excellence in Cardiac Care, I was one of the project directors for that program, that when you use quality improvement you can actually use it as a non-controversial solution to eliminating disparities because, as people focus on getting as close to 100 percent compliance with the quality indicators as possible, any disparity that existed prior, it shrinks, it goes away, it vanishes just by definition because people aren't so focused on all the qualitative and emotional aspects of disparities but they're focusing on the right thing, and that is providing the high quality evidence-based medicine to every single patient regardless of their background. And when they do that and they start achieving really high compliance rates with the quality indicators, their disparities disappear. And we saw that in that project, and so I'm an advocate for

using quality improvement to actually eradicate disparities.

CO-CHAIR GIBBONS: Thank you for bringing that up. I'd actually forgotten about that. I was present at the presentation of that at the national NQF meeting October 2009 perhaps, maybe 2008, one of those wonderful years, and I can remember that graph which was pretty darn clear with respect to the gap narrowing. It was very, very impressive. Thank you for bringing that up.

Okay. So, Helen, do you need any more guidance about this issue? Okay, good. Now, should we look at the portfolio?

DR. WINKLER: A couple sort of big picture questions that go along with your role as providing guidance for evaluating the portfolio. The first one is three measure developers have asked for measures that were previously endorsed by NQF to be retired, and mostly these are not

supported anymore. The question we would just ask you is do you have any particular feedback, questions, concerns? Does it give you heartburn that these are going away? We just want to take advantage of your expertise and the fact that you're kind of looking in this context, as we, you know, kind of look at the whole portfolio.

The measures are, the first one is 72. This is beta blocker treatment after heart attack from NCQA. This is a HEDIS measure that they actually stopped using because performance rates became very, very high.

The second one is an AMI inpatient mortality. This is the measure originally from the Joint Commission. The Joint Commission has taken that out of use in favor of the 30-day mortality that they work with jointly with CMS. So that measure is, I think they've still got it in their QI portfolio, but they're asking it be retired

from NQF's portfolio.

And the last one is from ACC, and one of the very earliest measures endorsed by NQF back in 2002 or something like that was PCI volume, and that measure has been asked to be retired. So --

CO-CHAIR GIBBONS: So why don't we take these in order. The first one, the beta blocker, Rochelle, you were the primary reviewer on the PCPI measure related to this. Thoughts about this?

DR. AYALA: Now, is this the one --

CO-CHAIR GIBBONS: This is not the measure you reviewed. This is a potential competing measure that's being retired.

DR. AYALA: Now, is this one in the hospital or when they get discharged?

DR. WINKLER: No, this is a, this is a health plan level measure, and it is looking at patients who have had a heart

attack but in the outpatient realm. So they're looking at usually prescription data.

DR. RASMUSSEN: So this measure was actually a beta blocker prescription within seven days of discharge, so, again, one of those single point adherence measures. And now the measure we discussed yesterday is that more of a long-term adherence measure. So really peg the needle on this one that we're talking about and now extend it out to 180 days medication adherence.

DR. RUSSO: The only comment, as we retire them, and it certainly makes sense to do so, and, obviously, I like beta blockers, but if you're retiring all of the beta blocker measures I guess the only thing I would wonder is, number one, should we look back in a few years to make sure that now people don't focus on something else and that slips off and make that a standard, or

do we just promote when we retire, if we retire all the beta blocker measures, do we say we have a composite measure that somehow we can get at it in the future or some other long-term beta blocker treatment that may be more important than initial treatment.

DR. WINKLER: One of the issues with these measures is they're no longer supported by their measure steward, so they won't be available for future use. That's different than, you know, keeping them out of NQF's portfolio.

DR. AYALA: I just want to clarify that we retired the beta blocker prescription at discharge but not the one for persistence of beta blocker use after --

DR. WINKLER: Well, just to clarify, you haven't done anything final yet, all right? That's why we've got the rest of this agenda. You've made initial steps, as I've described in the beginning, step one, the initial evaluation. We still

have a bunch of work to do before your final recommendations, okay?

CO-CHAIR GIBBONS: This is to make sure you stay awake for the remainder of the meeting. Other questions about the beta blockers? I think we feel we're in a better place with the measure we have. That's the sense. Okay. AMI inpatient mortality. Tom, do you want to speak to this since you reviewed the 30-day one? Retiring this one? Do you have any angst over this?

DR. KOTTKE: No, I guess not. There have been arguments that there are important lessons from inpatient mortality, but I think 30-day mortality is probably more consistent with the systems that are needed for good outcomes for patients.

DR. WINKLER: Just to mention, trying to keep things simple but it does get complicated, is currently there is another inpatient measure from AHRQ on inpatient

hospital mortality that is at the very end of the endorsement process. We don't lose it all together.

CO-CHAIR GIBBONS: Any other concerns about retiring inpatient mortality? I'm sorry. The Joint Commission's version of inpatient mortality. All right. And PCI volume, Sid, do you want to comment on that and any of the other interventionists?

DR. SMITH: Well, I'm not sure who decided to retire the PCI volume.

CO-CHAIR GIBBONS: The ACC.

DR. SMITH: So I haven't seen the thinking behind that. Initially, the concern to which PCI volume addressed itself was our need to identify quality of outcomes, and there had been some data to suggest that the more frequently the procedure was performed the more likely it was to be performed well. And volume criteria had been submitted both for hospitals and for operators. In addition,

there had been some data to suggest minimal volumes for hospitals and operators that were performing PCI for STEMI, and those had been written into the guidelines with the clear statement that they were only a partial method to really assess outcomes and that, when other approaches to assessing outcomes were available, that the centers were encouraged to participate in them. And along the way there have been many, many criticisms. It's very difficult with low volume to really determine quality. The fact that people are doing a lot of PCIs doesn't necessarily mean that they are doing them well and, in fact, there's been one really notable problem in California where angioplasties were being done for the wrong reason. And then we have issues about people starting up, issues about people that are injured while they're skiing, and 75 year-old medical reasons.

So there are a lot of problems

and there's been a lot of controversy about this. I think that the reason that it's being recommended that this be retired was what we talked about today. We have a new program here to look at mortality associated with PCI, and I think that probably is being retired because it's felt that there are better ways to assess outcomes than just look at the number of procedures.

In many hospitals now, that has been done by the hospital committees themselves. So speaking from a level of ignorance about why it's being retired, that's why I would think it is, and I think it's a reasonable suggestion.

CO-CHAIR GIBBONS: David?

DR. MAGID: Yes. I published two articles, and there's one in New England Journal and one in JAMA, so I think --

DR. SMITH: Yes, I've quoted your articles, and we've used them. They're good.

DR. MAGID: Thanks. I would say that, first, so we looked at institutional volume, not provider volume, and the threshold is fairly low. And my guess is is that, and Fred can speak to this, that the number of institutions that are below that threshold that report to the ACC NCDR PCI is very, very small. And then the other thing is, I agree, I think mortality is a better outcome measure because it's obviously the ultimate outcome that we're interested in, so I would say you can probably get rid of it.

DR. SMITH: Yes. I think everybody has wanted to move beyond a number of procedures to a better assessment of quality, and we've seen examples here that we're on the way to that.

DR. WINKLER: I think this is just a reflection of how things have evolved over time. And as better measures come along, it can measure more important robust

aspects of care, such as outcomes. Some measures have just outlived their usefulness. So thank you. We wanted your, I don't know, reactions to those before we recommend them to the Board that they are permanently removed from the portfolio.

Okay. A couple of other follow-up things. I think, at this point, I just want to mention to you what the next steps are before we talk about just how we're going to look at the whole portfolio. But during the course of your conversations over the last two days you've raised a lot of questions that were either partially or not totally answered, to say nothing of the fact you've raised the disparities to the top of that list of questions. What we're going to be doing over the next few days is preparing a series of questions to go back to the measure developer to get responses for you.

This is an iterative process. It's a dialogue. They gave you the

submission, you've responded to it. We're going to go back. We're going to volley a few times so that, as we move through time, you will have greater understanding of what the measures' strengths and weaknesses are. So just be aware that that's an ongoing process.

Once a steering committee has met like this, our dialogues with the measure developers are sort of on a very frequent basis. We become the very best of friends and regular e-mail buddies. So just realize that that's really what's going to go on. This does not stop. It really is part of a fluid ongoing process. So we will be bringing back some of these responses.

In that, some of the answers and some of the information that may be added to the mix that you haven't seen before will help in terms of the final resolutions. Probably the biggest question of information we're going to need to approach the

developers about is this issue of harmonization. In addition to harmonization, you all have identified and raised issues of measures that are so similar that they're really competing. You know, they're measuring the same thing. Is there a point of having two, or does that really become every confusing out in the world?

So that is a huge task, particularly with this set of measures, so we did not want you to really get embroiled in that at this setting. Now that you've had your first pass through the measures, we will want to be going back through that.

Let me just show you something that I have not yet shared with you but we will get there. And this is an additional sort of embellished spreadsheet. We've shared this with you at the beginning. This is the portfolio of measures. They're organized along the episode of care

framework. The measures that are highlighted in yellow were the ones that you looked at. They were either from a review or new submissions.

What we've also done is begun looking at them in terms of the measures that are competing or require harmonization. If you noticed, as we've scroll through, it is not a short list. And you've raised this issue many times. One of the more complex issues about harmonization is sometimes we need to harmonize the numerator with a group of measures and then the denominator with a different group of measures. The inclusions, the coding for things like AMI, CAD, ischemic vascular disease are amazingly off by two or three or five little inclusions and codes when you put them side by side.

So we will be doing these multiple side-by-sides, and there are not just one or two. There are a large number

of them.

The first thing we're going to do for harmonization purposes is we will be taking it back to the measure developers and get them together and say, look, you both are trying to measure ischemic vascular disease, you code it this way and you code it this way; what is the deal here, you know, why aren't they the same, and see if we can get that harmonization to occur. It is really their job to do. It's your job to reflect and evaluate how well they did it and whether they've actually done it well.

When it comes to actual competing measures, this is an area that has become very much a current topic. In one of our projects that's recently trying to come to a conclusion, the Board of Directors has sort of pushed back on us and said, look, these sound so much alike, it sounds like they're competing, you need to help us, you know, understand this issue of very, very similar

measures. As a result, policy is evolving as we speak, and Helen will be happy to share with you a decision tree that is almost final that we will be using.

DR. BURSTIN: So this actually is pretty hard to see. We'll make sure you get this individually, but we've been working and literally have a call next week with our board to finalize this, some guidance on competing measures. And so, essentially, trying to define clearly what is and what is not a competing measure. So same measure focus, i.e. target process, condition, event, or outcome, and same target population. So you want to just be really clear on what we're talking about in terms of which ones are actually competing.

We then go through a process where we say if they're competing measures how would you assess superiority? And that's going to be the next step here after we get some clarity on this. So, for

example, importance to measure and report, probably not going to be much of a difference there as based on the evidence, probably the same gap if they're competing measures, etcetera. So it really comes down to the next set of them. So for scientific acceptability, for example, untested measures of which you don't have, I think, very, very few, cannot be considered superior to tested measures, for example, as one point we put forward.

We would also ask you to look in terms of the specifications and the methods. Can you pick the measure with the broadest possible applications, settings, target populations, compare on reliability and validity, if you have that data. And then usability, all else being equal, if something is being publically reported it's preferred. If a measure has got the widest use, settings, number of entities, etcetera, it's preferred. And measures that are in

use are preferred over those that are just newly done and have never put out there at all.

And then on feasibility, again, measures of electronic sources, of course, given the feasibility concerns, are preferred. And measures that are freely available are preferred, as well.

But, finally, we recognize, even if you go through all that, and that's the situation we're in with the three measures we're going back to the Board on next week, sometimes you're going to come to the point where they're competing measures and there's no clear superiority based on those criteria. And so there we've been trying to come up with some guidance to say if there's not clear superiority can you justify having endorsement of multiple measures? And does that added value offset any negative impact?

So, for example, if perhaps the measure allows you to move more easily

towards and EHR-based measure, is that something to consider? Or if the additional measure is applicable to an additional setting or significantly increases the number of entities that you could capture, is that another reason to do it? But then the key thing there would be those two measures that have to be harmonized.

So we'll bring this to you in final form as you go through your next process, but we want to, you know, going back to Dr. Smith's comment earlier, we're trying to give you as much guidance as I think you're going to need to go into this brave new world for us in this next phase of work. So more to follow, but if you have any thoughts please let us know. And we'll share this with you on e-mail.

CO-CHAIR GIBBONS: So, Helen, can we put this little flow diagram to a little test?

DR. BURSTIN: Sure.

CO-CHAIR GIBBONS: Do you think it's ready for that?

DR. BURSTIN: It's almost cooked.

CO-CHAIR GIBBONS: All right. So we had this discussion earlier today about the median time to fibrinolysis and fibrinolytic therapy received within 30 minutes. The percentage versus the median. Can we sort of help me, those who discussed these, to try to apply these criteria to that situation.

DR. MAGID: Can I just ask a question about that? I thought that the issue was it wasn't extra work to provide the --

CO-CHAIR GIBBONS: No, wait, wait, I know. But we're just trying to apply, these are the criteria --

DR. MAGID: No, no, but I know. But I thought we heard that there were constituents who preferred it one way versus another, and if there's not any extra work

is that a problem to present it more than one way?

MS. PACE: Right. But I think that may ultimately be your conclusion, but we're starting with the idea they're both trying to measure the therapy given at the right time in the same population. So the question is, you know, they're trying to do the same thing and do we need both.

DR. AYALA: And one other question. So this is the issue at hand with things that we already have and even looking forward. So you're going to have new measures. We're assuming that people who developed them have looked at everything that's out there. Should there be something on the application that says if you see this might be a competing measure let us know the pluses and --

CO-CHAIR GIBBONS: It's there.

DR. AYALA: It is. So you have a lower chance of getting approved because

it's competing, but show us why it's better and what might be --

CO-CHAIR GIBBONS: But until this process plays out to actually, you know, identifying or choosing between competing measures, I think we saw in this round of applications that a number of applications just chose to politely ignore that blank. They really did not give much credibility, their answer wasn't credible. I'll say that. I suspect they just didn't think that blank was important.

MS. PACE: So we are on the next version of the measure submission, being even more directive about that. We're asking measure developers to attest that they've actually identified and worked on these harmonization and competing measures issues. Otherwise, it will not be accepted for consideration.

CO-CHAIR GIBBONS: I hear tough love works very well. Okay. So for those

in the back of the room, why don't you read the first box and then we'll try to apply it to this group because I think they're having trouble. To be honest, I'm having trouble with my bifocals in the front of the room.

DR. WINKLER: Well, the first one is does anybody disagree that those would be competing measures? They're measuring the same measure focus, same target population. So it brings us into importance to measure and report. You should have the same information on opportunity for improvement and the evidence base for the measure. So that should be a wash. So we move into scientific acceptability, and both of the measures are tested, so that does not give us any discrimination. And then we're looking at measures with the broadest application or comparison of reliability and validity on the overall criterion.

DR. KOPLAN: Sorry to interrupt. In this particular instance, the reason why

David maybe had a little issue with this example, if you go all the way to the top of the -- oh, sorry. Where it says numerator. The numerators are different, right? If we're just using this example, you would -- in other words, you would stop right there. That's why it might have been better to do, like as an example, use aspirin and MI versus aspirin and ischemic vascular disease because the numerators are the same and the other stuff is different.

DR. WINKLER: Yes, but I think what Karen would say to you is that the numerator, though, is you're measuring the same, you're measuring the same thing: time. So I don't know if it has to be that identical.

CO-CHAIR GIBBONS: If you just took numerator out of there and just said same measure focus -

DR. KOPLAN: But in this example, in this example, that is the only issue. So

the whole rest of the stuff, you are going to get to the end - so it didn't work.

CO-CHAIR GIBBONS: It would be worse if she was crying.

MS. PACE: Some people don't understand what we mean by measure focus because they're totally focused on terminology of the numerator, so that's the reason for the parens. But we can just as easily take it out. But the idea is we're not looking for measures that are exactly the same because if that's our criteria then we would not have any competing measures because they always differ by something. If they were exactly the same it would be the same measure. So we're really looking at measures conceptually first that are really trying to measure the same clinical phenomenon or condition or --

CO-CHAIR GIBBONS: Well, I'm going to keep driving this because you've got this draft. So if we come down and say

compare usability. Now, there's a statement that measures that a publically reported are preferred, but these are both publically reported, measures with the widest use are preferred, same, you know. So I don't know. So then we get down to -- yes.

MS. PACE: But I want to go back because I think that perhaps one of the things that we have to deal with and it needs to be discussed further, if we go back up to scientific acceptability. So they have different methods of getting at the same issue. And should we talk about is there some bullet point or some way to look at which is the more valid way to measure that? And then on the usability side, which one gives you better information on which to drive improvement? So we're probably missing some things, but that's what we'd like to -- I mean, how would you compare them? Do you think it's a methodology issue? Is there one that's a better

methodology in terms --

MS. SZUMANSKI: I think if you talk to people in the trenches, they're going to say that the percentiles that you see in median time to fibrinolysis are very usable to them, which may be different than the other measure. The question is who are you producing this information for? And I think that is the splitting difference between these two particular measures on the same topic.

CO-CHAIR GIBBONS: I think it really does boil down to that. I mean, I'm surprised they're both publically reported, to be quite honest. I'd love to see the survey of Americans as to what percentage of Americans know what the word "median" means. So from the public, you know, patient perspective, that one is virtually worthless. On the other hand, for quality improvement, that one is better than the point estimate or percentage below the

cutoff. So you have an argument both ways on this.

DR. MAGID: You know what would be interesting to see would be whether institutions really are ranked in a different quartile on the different measures. If you find that they're really ranked in the same quartile across both measures then they're really not providing any additional information, whereas something like mean, where you could have one outlier that pulls things up, it wouldn't be surprising to see a difference in ranking.

CO-CHAIR GIBBONS: Since we had an extended discussion about this one earlier today, while it's fresh in everybody's mind, I really think we should come to a decision on this one. That is, as we understood it, no incremental work of having a second measure, and we have one measure that's, I would argue, more patient

friendly and another measure that's more provider friendly with respect to quality improvement. How does everybody feel about continuing both measures? Does that seem reasonable, given that construct? And then we'd invite comments from the public.

DR. RUSSO: The only comment I would have, although it's no additional work to collect it, we are reviewing two separate measures each time. It's minimal additional work for us, I guess every three years, or is it even an option to say go back to the developer and say, hey, listen, can you combine these into one and measure "or" or "and' or one of the other in the same measure so you review just one measure each time, or can they pick? Maybe they have a preference.

DR. SNOW: They're so close to each other. They're two elements of the same concept. We just put them together, line one, line two, with one number, one

measure with two arms; is that wrong?

CO-CHAIR GIBBONS: Well, there are two different lines on Hospital Compare, I believe. Somebody can help me. Fred, is that right?

DR. KOPLAN: When people compare medians, I worry. I'm not as experienced with this kind of thing, but, you know, if there's some standard you have to meet where there's a cutoff, a percentage, like everybody is greater than 90 percent of X, of some standard, then hospitals can be equivalent when there's a median. If one has a number of 87 and another has a number of 86, and it's on Hospital Compare and people look at that, they're going to say, some people will say that one is better than the other one, even though the difference may not be significant. So don't you tend to usually always say, okay, if you meet this cutoff you get a passing grade and that's kind of a better way to approach

quality, or am I wrong about that?

MS. PACE: I think that there's differences of opinion about that. I mean, the phenomenon you're talking about also kind of gets into the reporting issue and whether you identify the amount of error around a point estimate to show that there's no difference between 91 and 90. You know, so some of this kind of gets over into how the data are displayed versus the actual measure construction, but it's a good point.

CO-CHAIR GIBBONS: There is an unforeseen kind of consequence which we have to, at least theoretically, consider, which is that, in the course of taking care of an individual patient, somebody realizes they're already past the threshold. They'll be less likely to hurry if they know they've already failed, whereas a median will capture that data so they'll presumably have an incentive to keep hurrying. I have no idea how often that happens, but I'm sure it

happens because everything happens.

DR. AYALA: In terms of operational, I'm wondering how the median impacts the quality improvement process more than the percentage. Because in the median process you can actually throw out your outliers, but those are, you know, those are included when you're looking at your percentage of compliance.

CO-CHAIR GIBBONS: Fred, do you want to try to answer that question?

DR. MASOUDI: I mean, these are just things we've heard in implementation. I can't tell you why people, you know, necessarily like these things but --

DR. AYALA: I'm just going to share my experience for a moment. I was actually asked, when I was doing the Robert Johnson Foundation project, to serve as the chairperson as the PCI task force, and every time we had a fallout, even if it was just like by one minute, oh, my gosh, you should

have seen the angst that everyone went through, and we had to look at every single time segment along the process. And we really took everything really seriously. My feeling about the median is that it might actually cause a bit of a more relaxed approach to the quality improvement process because then you can get rid of your outliers. And outliers are really important because, in our situation, the outliers tended to be the patients who came with atypical chest pain or no chest pain at all, you know, the atypical presentations where we had to cast our net wider to capture those early so we got that EKG within, actually ten minutes is a good estimate but less than ten minutes.

So, to me, the real true attention to the indicator itself, which is evidence based, to me, it's more pressing when it comes to your operational quality improvement process. That's just my

experience.

DR. BURSTIN: I just checked on Hospital Compare, and, at least currently, only the two threshold measures are reported, not the median. So one could make the argument this is probably reported and very useful for internal QI, so maybe these should continue to be the ones publically reported. So consideration for you.

CO-CHAIR GIBBONS: By your criteria then, the proportion would win. We've just gone through it. Everything else is equal, so the proportion would win because it's publically reported.

DR. KING: I have a question. Have you considered that this might be what we call a technicality? The measure is the time to get fibrinolytic therapy. You can express that as a time, a percent that gets it under a certain number of minutes. You can express it as the number of people 50 and over that get it. You can express it as

the number of males or females or blacks or whites or Hispanics that get it. Each of these measures, and there has been no restriction that I've heard thus far, has multiple ways of displaying. And we, just a few minutes ago, President Chairman, encouraged, indeed demanded that they use the measure and report it in more ways. So I say, sir, in the sake of harmony, that we, by fiat, declare that this is, in fact, one measure, it already is, and the consumers of it can see the data however they'd like.

CO-CHAIR GIBBONS: Well, that's going to be hard for anybody to say anything after that. So we'll just vote on that proposal, which would, in essence, leave both of them as is as different expressions of the same data regarding time to fibrinolysis. All in favor? Opposed? All right.

DR. BURSTIN: That's why I think coming back and actually presenting this to

you a little more thoughtfully with tables would be nice just because --

CO-CHAIR GIBBONS: Well, I think we had a volunteer to try a second example, if you're willing.

DR. BURSTIN: I think a second example, if you guys are willing, would be great. I'm just saying that the first one is a little complicated because one of the measures is only, I think, for transfer patients. So we just want to think about this. It's a little nuanced, though.

DR. MAGID: That's the issue with the aspirin, so ti's going to be hard to harmonize.

CO-CHAIR GIBBONS: Actually, no, both of those are on transfer patients, I believe.

DR. WINKLER: Yes. The pair is for the transfer measures, the percentage is the only one for hospital measures.

CO-CHAIR GIBBONS: Okay. So I

think the other eager volunteer to try this framework -- well, we want to test the framework -- occurred earlier in the day yesterday. Now, I know this has been a wonderful experience and you probably have difficulty remembering that, but we did consider vascular disease use of aspirin or anti-thrombotics and CAD antiplatelet therapy back to back, okay? And Bruce had the first one.

DR. KOPLAN: I remember mine said the CAD antiplatelet therapy, or mine was aspirin and vascular disease, and then was the person who did -- okay.

CO-CHAIR GIBBONS: George had the other one.

DR. KOPLAN: So in terms of going through this, the competing measures, so the measure focus appears to be the same. Would you agree, George?

DR. PHILIPPIDES: I think one was vascular disease --

DR. KOPLAN: Oh, that's right. Mine is --

CO-CHAIR GIBBONS: That's going to be the target population --

DR. KOPLAN: So mine is more encompassing, I shouldn't say mine. The aspirin in ischemic vascular disease. No, I don't want to be, I'm already responsible for all the right heart caths in the 1990s in Massachusetts. So ischemic vascular disease is a wider net than CAD. That's the denominator, and the numerators are the same.

CO-CHAIR GIBBONS: So, Helen, per this construct, does it end there? Because there were other major differences in these measures when we were discussing them.

MS. PACE: Well, one of the things that we didn't include here was actually a discussion before we get to the competing is whether something like that should be combined into one measure.

DR. BURSTIN: So for example, could it be IVD with a strata for CAD, if you think it's important enough to have CAD separate. Just an example of perhaps ways to approach that.

CO-CHAIR GIBBONS: Well, as I recall, one of the dilemmas of having those two back to back was that the exclusions were different and were much more carefully defined, as I recall, from a clinical standpoint in the second measure that George reviewed. Is that right, George?

DR. KOPLAN: Agree, yes. I remember that, too.

CO-CHAIR GIBBONS: Yes.

DR. KOPLAN: And people had the issues with the lack of exclusions in the aspirin in ischemic vascular disease.

CO-CHAIR GIBBONS: So I think that that experience would suggest there's a potential to broaden perhaps, I mean I think there was a solution for that one.

DR. KOPLAN: The decision tree is a tree you have to go down even if you stop at one branch. You still have to go down the tree anyway.

MS. PACE: I mean, because, you know, you point out a good thing. I mean, strictly speaking, those were different denominator populations, but the question is should they be? I mean, does the evidence indicate that aspirin is really indicated for the broader population? Then if there's some way to work through that, either combining or --

CO-CHAIR GIBBONS: So we're now starting assignments for the next interaction of the meeting. And you guys were which group?

DR. KOPLAN: Three.

CO-CHAIR GIBBONS: Three. So I would suggest, hearing no objections, that group three be tasked with looking at those two, 0068 and 0067, with respect to

suggestions for reconfiguring those into a single measure and then which framework, which measure developer gets that feedback and tasked with doing that. Because it would seem to me that that's an example of something where we could conceivably eliminate a measure by creating one very good one.

DR. RASMUSSEN: What are the implications if we have two measures that we harmonize that have different developers?

CO-CHAIR GIBBONS: If they can be harmonized using the same platform, in terms of criteria, I would suggest that it can go forward that way. But in this case, if we were to favor, for example, the exclusions listed in the AMA proposal, the NCQA process would not allow them, as described by their staff yesterday.

DR. RASMUSSEN: So to frame it even more specifically, there can be one owner and one developer. If we harmonize

two measures, one of the developers would have to give up ownership.

DR. BURSTIN: We actually do have some examples of co-ownership, but they would need to come together and agree. But it takes a long time I'll warn you, having just spent about five months trying to get one C-spine measure combined. It takes a long time.

MS. PACE: The other thing, and I think what Reva said is that for the ones that are clearly harmonization issues versus competing, you know, she's going to go back to the developer to ask them to harmonize. Now, when they're competing and you're trying to make a decision of one or the other, that's why if you can identify one that's clearly superior, that's the more efficient route because trying to get two developers, after they've invested in a particular measure, as Helen said, is quite lengthy.

CO-CHAIR GIBBONS: Okay. What group were you in? What group number? Four? So I would suggest a similar task for group four with respect to measures 0075 and 0074 on lipid control, which Mary discussed both of those yesterday. And as we went through those, I think there were discernable differences in the way they approached exclusions that potentially offer an opportunity for us to take a position. And I think it best done not on the spur of the moment here but after careful due deliberation with whatever kind of wine in hand you want for that evening.

DR. WINKLER: And side-by-side tables from us, so we'll make it easy for you to see the similarities and differences.

CO-CHAIR GEORGE: With these being both outpatient measures, do you see a need to try to harmonize with similar discharge measures from the hospital setting?

DR. WINKLER: I think we will ask you that question and, you know, add that into the mix. I think Mark was the one who noticed, what is it, five measures for aspirin use? So I do think that is a question for you to address. You may decide that it's okay to have some differences based on setting, but I think it's important that you consider it explicitly and be able to provide the rationale for that. So as I said, this is very complex how we're going to have to make these multiple comparisons.

MS. SZUMANSKI: I think there are two other variables that are not listed up here that we've mentioned today. One is on diversity, the impact on diversity, diverse populations. And the second one, as difficult as it is to think about, is the financial impact of the monitoring of that measure or what impact does it have on the institution in terms of their reimbursement, etcetera. So I think those are not listed

here and may or may not be important, but I think they're worth being said at least.

CO-CHAIR GIBBONS: I think they're more than worth being said. Thank you, Kathleen. I think they're both very important points for this process. And Helen is intensively revising the grid as we speak. It's just become two pages or else it's one page of impossible to read print under any circumstances.

MS. PACE: You actually didn't get the actual algorithm and other things.

CO-CHAIR GIBBONS: All right. We've talked about the antiplatelet issue. We've talked about the lipid issue. I think the other issue that repeatedly surfaced in various ways is blood pressure. And we've given a clear message back to the Minnesota Community Measurement Project and, hopefully, we'll see a revised submission from them, but the NQF has already endorsed a blood pressure measurement for diabetes,

blood pressure measure I should say for diabetes. DR. WINKLER: And you will see the measure for blood pressure control for hypertension in phase two.

CO-CHAIR GIBBONS: So what I would respectfully suggest here is that everybody put their thinking caps on because we can't have potpourri of confusion for the remainder of 2011, pending the release of JNC 8. And I do think we want to take a position that seeks to have a uniform blood pressure standard. And this is another big topic but I would like people to think about it. I would throw on the table a strawman which is that all of these developers should be told in very clear terms that they have to comply with JNC 8 pronto to avoid confusion in the practice community because I think that's one thing that drives docs nuts is to see different "guidelines" from different groups, and blood pressure, it seems to me, should be driven by the

national process that NHLBI has directed for years. Mark?

DR. SANZ: So having said that, how quickly and what's the mechanism so that 15 busy people don't have to come together to approve some change in the measure?

CO-CHAIR GIBBONS: Staff, help.

DR. WINKLER: Well, Mark, what are you asking? In terms of what level?

DR. SANZ: Well, I think the guideline is --

DR. WINKLER: Right, okay. All right. Whenever evidence changes, major guideline changes, NQF has a process of having an expeditious ad hoc review that doesn't require the entire world to come together and talk about it that we would put into play for these measures.

CO-CHAIR GIBBONS: We were thinking about a simple site visit to Montana to discuss it.

DR. WINKLER: But this is not an

infrequent thing. If you recall the ACCORD trial, I mean, lots of things happen on a regular basis, so we have had to deal with this issue previously. It's not a new problem.

CO-CHAIR GIBBONS: Are there other issues of harmonization that were mentioned as we went through, particularly now for the primary reviewers, remembering that blank, that we need more extensive kind of prep for for our next meeting? Those are the three that I identified as we went through. Mark?

DR. SANZ: I don't know if it's another issue, but maybe you covered it and I just don't remember. Why does it start, like 67 and 68 are two different organizations, why can't they be asked in the next 30 days to review their own criteria before we come back in April and say what they can or cannot accomplish rather than we have to do it for them? I

mean, we have other things to do.

DR. WINKLER: Mark, I'm sorry if I wasn't clear, but that's exactly what I said we were going to do.

CO-CHAIR GIBBONS: But I do think it's worth pointing out that, as Helen has politely indicated, this is a sometimes difficult and long process. So she is still dealing with the directives from the last committee I served on when I had more hair, so you can tell that it's taken a while for this to play out.

DR. WINKLER: One of the things that, having listened to you over the last couple of days, for our meeting in April, which I'm going to remind you we're going to be looking at an additional 23 measures and looking at a different topic area, this is sort of the etcetera group, the hypertension, atrial fib, heart failure that fall into this cardiovascular bucket. There are only 23 measures. Most of them are

maintenance measures. But in preparation for that, we are going to go back to the developers who've submitted their measures and first we're going to go back and review their submissions for things like information on disparities and some of these other questions you guys have raised. And if it looks like they really have not submitted appropriate information, we'll go back to them and say, you know, it would be in your best interest to fill in the blank because the committee is not going to see this favorably with no information. So it will be an opportunity for them, as opposed to a requirement. But we can certainly do that for our April meeting.

We will need to get back with you to finish the work on these measures. We are going to have to schedule a couple of conference calls. Hopefully, we can do a bunch of this by e-mail. But as you can see, this is a complex task as we try and

sort through all of these various issues, particularly the competing measures and harmonization issues. This is the first project where we've had this level of so many measures being involved in the need for harmonization. Usually, it's a one or two kind of thing, not every measure you've looked at practically. So that provides its own sets of new challenges.

CO-CHAIR GIBBONS: All right. So the other thing that we want to briefly deal with is the question of gaps in measures that adequately describe the clinical care process. This deals with, if you remember, I don't know how to describe that diagram across the board with the different process steps. We had that at one point for an imaging conference, and we all called it the Masoudi diagram because Fred drew it on an envelope and we ended up using it in the publication. But in any case, that's sort of, whatever that's called -- DR. WINKLER:

The bubble diagram.

CO-CHAIR GIBBONS: The bubble diagram. And I would suggest that there were actually examples that several of you cited during these discussions, and so this is the time to sort of put them on the table for staff to mull them over and put them into some sort of comprehensive form for the developers. So one, for example, is the point that Dana raised, which was do we have a measure for people not going in the hospital that reflects, basically, the goal of good outpatient care, which is to keep people from going in the hospital. That's actually part of the continuum of care at this point in time, as we heard from, I think it was from, AHRQ, quote, could be measured but isn't being measured.

The other one was the universe of hospitals, which Sid kept us pointed towards because he kept counting up the hospitals and trying to figure out what happened to

them all. And I think that alludes to sort of the whole issue of what is happening in those, be it hospitals, practices, whatever, who aren't participating in the voluntary submission of data or the various registries that are the sources of some of these proposals.

But I'm sure some of you thought about this as you looked at your individual measure and thought about what needs to be improved. So are there ideas that you want to offer at this time, this is just free-floating to sort of get them on the table while they're fresh in your mind of things where measures might really be needed. It isn't to say there's a data source right now, it isn't to say there's a track record, but just to say this is something worth doing. So I throw it open for ideas.

DR. RUSSO: Is it even a possibility to even consider, so people don't just pick a measure, I mean they

obviously can pick measures, but if there's a group of measures they need to pick all three as opposed to picking one that might prevent some cherry-picking of measures that they know they'll perform well in? I don't know if that's --

DR. WINKLER: One of the techniques that NQF has used through the years has been pairing of measures, and pairing is often a bad word when we're talking about more than two, but grouping measures such that the recommendation with the endorsement is you don't use one, you use all of them. So a paired set is they're paired for use as endorsed measures and you do them both. We have groups of three and ten and whatever, if necessary.

The one thing I heard you all say was use of composites. There seem to be a couple of opportunities for composites, say AMI discharge medications or PCI discharge medications, all of that. And you also

seemed to like the idea of some of the all-or-none composite approaches, did that individual patient get the three medications they were supposed to, as a way of continuing to promote those processes of care, even though right now the current individual measures are kind of pretty much topped out and unlikely to promote a whole lot more improvement as is. So things like that.

CO-CHAIR GIBBONS: So is that really what you were thinking of, Andrea?

DR. RUSSO: It would be easier than to have to create a whole composite measure and put more work into it. If we have the separate measures, then we'd have to -- it's not exactly the same thing, obviously, but at least you're going to require people to report. You can't pick something you didn't do well in and not report that.

CO-CHAIR GIBBONS: The one

advantage of that is you can then turn it into an all or none where you say, okay, you'll get a score if you do all three of these things. The IOM encouraged that in their original report on performance measures, and it's a hard bar then for everybody, so you take these very high adherence rates and suddenly they don't look so high because the experience in the state of Minnesota was that failure on one is poorly predictive of failure on a second one. Everybody thinks, oh, it must be the same doc or the same patient or whatever. Actually, it's not. They're almost mutually exclusive. So it's a kind of interesting thing where you're 92, 92, 92, 92, you'll actually come down well below 80 on the all or none. So there's some utility in doing that, and I think, as I cited the example which was repeatedly cited in the state of Minnesota when this was being proposed, we don't think you're delivering good care if

everything is perfect except the blood pressure and that's 220 over 120. That was a hard argument for any doc to counter.

Sid, you had a point.

DR. SMITH: Well, just a comment. We're seeing that in China. In Dongbei, they don't use the ACE inhibitors. In Sichuan, they don't use beta blockers. You get this regional variation in China, just like what we're seeing in the United States in terms of therapies. You can't say if one thing is not used uniformly the others will not be well. It seems to be sort of a heterogeneous situation.

CO-CHAIR GIBBONS: Yes, yes. Okay. Well, if any other thoughts come to you on the plane where I know you're going to be thinking more about this meeting, please jot them down and get them to us by e-mail. Next steps will be staff communicating with us by e-mail and probably set up a conference call before our meeting

in April. I know people are emptying out, but I wanted to thank everybody for your participation.

Oh, public comments. It looks like the public is also emptying out. And I just wanted to indicate that the Chair recognizes this as a quality improvement process, so any comments or suggestions you have for me feel free to e-mail or phone me. I'd welcome them. I'm trying to make this a good use of your time and, hopefully, a stimulating few days. So thanks again.

(Whereupon, the foregoing matter was concluded at 2:59 p.m.)

C

C E R T I F I C A T E

This is to certify that the foregoing transcript

In the matter of: Cardiovascular Endorsement Maintenance Steering Committee

Before: National Quality Forum

Date: 02-16-11

Place: Washington, DC

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