NATIONAL QUALITY FORUM

+ + + + +

SURGERY ENDORSEMENT MAINTENANCE 2010
STEERING COMMITTEE

+ + + + +

WEDNESDAY
MAY 4, 2011

+ + + + +

The Steering Committee convened, in the A. Philip Randolph Room at the Embassy Suites DC Convention Center, at 900 Tenth Street, Northwest, at 9:00 a.m., Arden M. Morris and David Torchiana, Co-Chairs, presiding.

PRESENT:

ARDEN MORRIS, Co-Chair, MD, MPH, FACS
DAVID TORCHIANA, Co-Chair, MD
NASIM AFSAR-MANESH, MD
HOWARD BARNEBEY, MD

JAMES E. CARPENTER, MD
ROBERT R. CIMA, MD, MA, FACS, FASCRS
CURTIS D. COLLINS, PharmD, MS, BCPS AQ-ID
PETER W. DILLON, MD, MSc
RICHARD P. DUTTON, MD, MBA
STEVEN FINDLAY, MPH

PAULA R. GRALING, DNP, RN, CNS, CNOR
VIVIENNE HALPERN, MD, FACS
EILEEN KENNEDY, *
RUTH KLEINPELL, PhD, RN, FAAN
JOHN MORTON, MD, MPH, FACS
DENNIS W. RIVENBURGH, MS, ATC, PA-C
TERRY ROGERS, MD

CHRISTOPHER SAIGAL, MD, MPH, FACS
NICHOLAS J. SEARS, MD
ALLAN SIPERSTEIN, MD

PRESENT (CONT.)

RENAE STAFFORD, MD, MPH, FACS

CAROL WILHOIT, MD, MS

CHRISTINE S. ZAMBRICKI, CRNA, MS, FAAN

NQF STAFF PRESENT:

HEIDI BOSSLEY

HELEN BURSTIN, MD, MPH

ALEXIS FORMAN

MELINDA MURPHY, RN, MS

JESSICA WEBER

ALSO PRESENT:

SKIP ANDERSON

SUSANNAH BERNHEIM

JOHN BOTT

DALE BRATZLER *

LAURA EATON *

SUSAN FITZGERALD

JEFFREY GEPPERT *

LAURA GROSSO *

JANE HAN *

WANDA JOHNSON *

TIM KRESOWIK *

VICTORIA LYNCH *

KRISTYNE McGUINN

PATRICK ROMANO

DAVID SHAHIAN

DAVID SHAPIRO

JEFFREY SILBER *

DONNA SLOSBURG

SUSAN WHITE *

* Present via telephone

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

C-O-N-T-E-N-T-S (CONTINUED)

C-O-N-T-E-N-T-S (CONTINUED)

C-O-N-T-E-N-T-S (CONTINUED)

C-O-N-T-E-N-T-S (CONTINUED)

C-O-N-T-E-N-T-S (CONTINUED)

C-O-N-T-E-N-T-S (CONTINUED)

C-O-N-T-E-N-T-S (CONTINUED)

C-O-N-T-E-N-T-S (CONTINUED)

C-O-N-T-E-N-T-S (CONTINUED)

C-O-N-T-E-N-T-S (CONTINUED)

C-O-N-T-E-N-T-S (CONTINUED)

C-O-N-T-E-N-T-S (CONTINUED)

C-O-N-T-E-N-T-S (CONTINUED)

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

9:00 a.m.

CO-CHAIR MORRIS: Good morning.

Hi, everybody. Thanks so much, again, for participating in the Steering Committee. It's nice to see you all here today.

What we are going to do right now is just, basically, welcome and introduction, very similar to last time. And I would like to just go around the table and have everybody introduce themselves, say where you are from. Then, we also need to verbally state whether we have any disclosures of conflict of interest or bias.

I will start. I am Arden Morris. I am the Chief of Colorectal Surgery at the University of Michigan, and I have no conflicts of interest.

CO-CHAIR TORCHIANA: I am David Torchiana.

I am happy to be here today, and I

want to add my thanks to all for coming.

I am formerly a cardiac surgeon. I was Chief of Cardiac Surgery at Mass General, and I am now the head of the physicians' group there. No conflicts.

MS. FORMAN: Alexis Forman, project staff.

MEMBER AFSAR-MANESH: Nasim Afsar. I am an internist and with the Department of Neurosurgery and the Associate Medical Director of Quality and Safety at UCLA Medical Center. I have no conflicts.

MEMBER SAIGAL: Chris Saigal. I am a urologist at UCLA. No conflicts.

MEMBER HALPERN: Vivienne Halpern, Chief of Vascular Surgery at the Carl T. Hayden Phoenix VA Medical Center. No conflicts.

MEMBER CIMA: Bob Cima. I am a colorectal surgeon and Vice Chair of the Department of Surgery for Quality at the Mayo Clinic, Rochester. I have no conflicts.

MEMBER WILHOIT: Carol Wilhoit, Quality Improvement Medical Director for Blue Cross/Blue Shield of Illinois. And I have no conflicts.

MEMBER BARNEBEY: I'm Howard Barnebey, an ophthalmologist from Seattle. I have just joined the Committee a week ago. So, some understanding and sympathy will be appreciated.

(Laughter.)

And I have no conflict of interest to report.

MEMBER GRALING: Good morning.

I am Paula Graling. I am the Clinical Nurse Specialist of Perioperative Services here in town at INOVA Fairfax Hospital, and I have no conflict.

MEMBER RIVENBURGH: Dennis Rivenburgh. I am a physician assistant with St. Anthony's Primary Care and Co-Chair of the Quality and Safety Committee of St. Anthony's Hospital. And I have no conflicts of

interest.

MEMBER COLLINS: Hi. Good morning.

Curtis Collins, Clinical Pharmacist, Infectious Diseases, at the University of Michigan. No conflicts.

MEMBER ZAMBRICKI: Hi. I'm Christine Zambricki. Good morning. I am a nurse anesthetist and I am Deputy Executive Director of the American Association of Nurse Anesthetists.

And I reported a potential conflict when I agreed to participate. That is that I sit on an advisory board to a surgical information systems company.

MEMBER STAFFORD: Hi. I am Renae Stafford. I am the Director of Surgical Critical Care at the University of North Carolina in Chapel Hill.

The only potential conflict is I am a member of the Quality Committee at SECM, and I am also a member of the American College

of Surgeons.

MEMBER KLEINPELL: Good morning.

Ruth Kleinpell. I am a Professor of Nursing at Rush University Medical Center in Chicago and also a nurse practitioner. And I have no conflicts.

MEMBER SIPERSTEIN: Allan Siperstein. I chair Endocrine Surgery at the Cleveland Clinic. I have no conflicts.

MEMBER ROGERS: Terry Rogers, CEO of the Foundation for Health Care Quality in Seattle. And I have no conflicts.

MEMBER CARPENTER: Good morning.

Jim Carpenter. I am the Chair of Orthopedic Surgery at the University of Michigan. I have no conflicts.

MEMBER FINDLAY: Hi. Good morning.

I am Steve Findlay. I am a Senior Health Policy Analyst at Consumers Union. I am a consumer representative on this panel, and I have no conflicts to report.

MEMBER DUTTON: Rick Dutton. I am a former trauma anesthesiologist in Baltimore, now the Executive Director of the Anesthesia Quality Institute in Chicago. I have no conflicts.

MEMBER DILLON: Peter Dillon, Chair of Surgery at Penn State, and I have no conflicts.

MS. BOSSLEY: I have two microphones. I don't know what to do.

(Laughter.)

Heidi Bossley, Vice President of Performance Measures at NQF.

MS. WEBER: Jessica Weber, NQF.

MS. MURPHY: Melinda Murphy, NQF.

CO-CHAIR MORRIS: All right. Thank you, everyone.

We are just going to next talk about our expectations and the process for the meeting. This will be very familiar for most of you, maybe not all of you.

Basically, our plan here is to

start, as you will see on your agendas, by talking about initially a recap of the work to date, which Alexis and Melinda will lead.

We will talk about how to use our remote controls appropriately to vote.

And we will talk about the Phase I measure followup. There were a lot of things about the Phase I measures that we wanted to get more information about, and there are some votes that we postponed until we received more information. So, that has happened in the interim.

We will start by talking about Measure 0300 and get the additional information from CMS that we were looking for last time.

And, then, we will determine whether or not our conditions are satisfied. We won't vote on this measure quite yet.

We will, then, go to the list on page 2 of the agenda and sequentially vote on those measures. We will not represent the

measures. So, those of you who are responsible for each of these individual measures won't necessarily need to represent unless the group feels that it would be beneficial prior to voting.

As you can see, our agenda is really full for today. We will try to be efficient with time, but, on the other hand, the reason that we are all here in person is to have a discussion. So, if we need to do that, we certainly will.

We also have -- is it time to open the meeting? Okay. So, it is time to open the meeting to the public.

Before we do, Dr. Morton, would you just introduce yourself to the group?

MEMBER MORTON: I am John Morton. I am an Associate Professor from Stanford University.

CO-CHAIR MORRIS: Thank you.

So, we are going to open it to the public now, and we have several folks on the

line and in person. So, we would like to ask you to introduce yourselves.

MS. SLOSBURG: Good morning.

I am Donna Slosburg with ASC Quality Collaboration. I guess I should have told you I am a measure developer.

MS. MURPHY: And those on the phone, anyone on the phone?

MR. BRATZLER: Dale Bratzler.

MS. MURPHY: I am sorry. Say your name again, please. It did not come through well.

MR. BRATZLER: Dale Bratzler.

MS. MURPHY: Hi, Dale.

Anyone else on the phone?

(No response.)

CO-CHAIR MORRIS: Let's see, I think that that's about it, and it is time to recap. So, already we are ahead of schedule. This is great.

(Laughter.)

MS. MURPHY: Okay. So, for the

recap, what we want to do is just quickly repeat some of the information from the last meeting. And we will start with the project purpose.

Again, when you are looking at the measures, you are looking at endorsement or continuation of measures for quality improvement and public reporting, and all these measures are around surgical patient and surgical procedures.

As part of this, there is a maintenance review of the surgery standards that were endorsed prior to June of 2008. There are some specific things that you looked at last time, asked for last time, that you will want to do again during Phase II. We will point those out as we look at the criteria.

So, for the Phase I measures, the first part of the agenda is to resolve any outstanding issues. You will recall that there were some requests that you made at the

time of the last in-person meeting that were resolved. The majority of them were resolved to your satisfaction during the March 31st call, and there is one that we will hear more about today.

Then, you will be making recommendations related to related and competing measures that you considered in Phase I and make endorsement recommendations for the measures, the 20 measures that you considered in Phase I that you determined met the four criteria that would make them eligible for consideration for endorsement.

And, then, we will move to the Phase II measures, where you will be evaluating 38 measures against the measure evaluation criteria, as you did for the Phase I measures. And you will, again, make recommendations related to related and competing measures.

And what we talked about a bit earlier this morning is that you might not be

ready to make final recommendations about related and competing measures, particularly Phase II, depending on where we are in the work by that point in the day tomorrow and, again, to make endorsement recommendations.

So, I want to just quickly run through the criteria.

Importance to measure and report is the first criteria. And, remember, these are in a hierarchical fashion, and you make determinations about importance before you move to any of the others, with the understanding that, if it is not important to measure and report, you probably don't want to even look at the other pieces with the exception of those measures that you might determine are topped-out.

And for those measures, if you determine that they are not important to be continued for endorsement because they are topped-out, you will want to look at the other criteria in order to make a determination

about whether or not they should go into an inactive status. You would not want to move a topped-out measure into an inactive status where it could be reactivated if you did not feel that it met all the other criteria.

So, for those measures, once you make the determination it's not important, and it is not important based solely on the fact that it is topped-out, we will still go through the other criteria.

And for the measures that are undergoing maintenance, a couple of the things that you will be looking for, in particular, is whether or not they have identified what the performance gap was and, within the performance gap, have they stratified it so you can see where disparities may be that might need to be further addressed? So, you are going to be looking for data, for numbers, as well as the citations.

In scientific acceptability, there is the list of things you are looking for with

respect to scientific acceptability with the fundamental piece being, are the measures reliable; are they valid? So, within that, you are going to look at the testing that has been done for reliability and validity, and, again, looking at the other pieces of the information around scientific acceptability.

For usability, if the extent to which the audience that is to receive the output of the measure can understand the results and will find them useful, and some of you on the Work Groups will recall that there were some discussions about some of the measures in terms of what is the utility to the public and how would you convey that information to the public. So, those are questions that you will answer as you think about usability.

Also looking at, are the measures harmonized? And again, whenever we are looking at related and competing measures, what do new measures add in terms of value to

the measures that are already endorsed and availability?

And with the maintenance review considerations, one important piece is, are the measures that have been endorsed over time being publicly reported? And if not, what are the plans for public reporting, just as you did in Phase I?

Feasibility. Is the information, the data required to compute the measure available? Can it be retrieved with relative ease, not a lot of additional burden? And can it be implemented?

So, you are looking for information, data that is generated as part of the process, the care process, information that is available electronically. If there are exclusions, do you have to go someplace other than the data source for the measure elements proper to get the information about the exclusion? If so, that is an additional burden.

What are the susceptibilities to inaccuracies? And given all of the information about where the data comes from, can the strategy that is recommended for use in collecting it actually be implemented?

So, then, in the course of rating those criteria, the first one, as I said, was importance. That is a yes/no rating.

The other criteria are evaluated based on the level to which they meet the criterion, either completely, partially, minimally, not at all, or it is not applicable. And you will get some more information as you look at how you are going to use the tool.

So, Jessica is going to give you information about the voting, but do you want to do that now or do you want to do it in a bit?

MS. WEBER: We can do it now.

MS. MURPHY: Okay.

MS. WEBER: So, everyone should

have the voting tool that they had last time. So, to vote, you press the number of your choice and then Send. And make sure you aim it towards me, so that it picks up the vote.

(Laughter.)

So, we can go ahead and do a demonstration, if you want. I will leave the measure up that we are going to vote on first, and you should have 60 seconds to put your vote in, once I start the timer.

It will be on the screen to my right, and it should give a live tally of how many votes we have.

(Whereupon, a voting demonstration was performed.)

We should have 19. Okay, great. All right. Thanks.

MS. MURPHY: Alexis?

MS. FORMAN: As Melinda and Dr. Morris indicated, we will start off with outstanding issues from Phase I. We will begin with a brief introduction from the

measure developer for Measure 0300.

The Committee suggested changing the measure from 6:00 a.m. to 24 hours. CMS had to go back to their expert panel to provide that suggestion. So, we will hear results from that meeting and if they are able to make their modification.

Then, following that discussion, as Dr. Morris indicated, we will make final recommendations on 15 of the Phase I measures. And unless someone indicates we need to discuss some of the measures, we will just go straight to the voting.

As Melinda indicated, we will then go into Phase-I-related and competing measures. And we will have a brief discussion on each of those measures.

This slide is just to indicate we pulled four measures from a current project that is going on, the Pediatric Cardiac Surgery Project. We brought them over for competing and related, which we will have a

discussion on tomorrow at 1:30 during our Phase-II-related and competing discussion.

These measures are related to two of our maintenance measures, 0339, pediatric heart surgery mortality, and 0340, pediatric heart surgery volume. So, we will discuss that in more detail tomorrow.

CO-CHAIR MORRIS: Great. So, this is the appropriate time, is it Dale, to --

MS. MURPHY: Yes.

CO-CHAIR MORRIS: Okay. For Dale Bratzler to go ahead and talk with us about the additional information regarding Measure 0300.

MR. BRATZLER: All right. Can you hear me okay?

CO-CHAIR MORRIS: It is actually real quiet. Is there a way to turn up the volume? Dale, we can barely hear you. So, we are going to work on the volume here.

MR. BRATZLER: All right. Let me know if you are having trouble.

Can you hear me?

CO-CHAIR MORRIS: Anybody who cannot hear him?

Okay. Dale, why don't you go ahead and proceed?

MR. BRATZLER: All right. We will get going, everyone.

After the last meeting, as you know, the Committee asked us to take the measure back and consider some changes to the measure. And the primary change that was requested at the time was that we look, rather than the previous performance metric, which the data element was collected 6:00 a.m. blood sugar or blood sugar that was collected closer to 6:00 a.m., that we collect the blood sugar at 24 hours.

We took the discussion back to two different groups. First, I think many of you are aware that the Society of Thoracic Surgeons has published a practice guideline series on blood glucose management during

adult cardiac surgery. It was published in the annals of Thoracic Surgery in 2009 and provides evidence-based recommendations for blood sugar control in cardiac surgery patients.

So, we had a conversation with several of the authors of that guideline. Then, we met with the Surgical Care Improvement Project Section Technical Expert Panel, who invited Dr. Tony Furnary, and I believe Dr. Richard Engelman was also on the call, both authors on the STS, to participate with the Technical Expert Panel.

So, let me just highlight just a couple of things. First, there is still strong support for the measure. We feel that, based on the conversation with the Technical Panel and STS, that there is strong evidence to support glucose control in cardiac surgery patients.

Secondly, when you look at the STS guideline, they make explicit Class 1A

recommendations around glucose control. I will just read the one recommendation that is relevant here.

"Patients with and without diabetes with persistently-elevated blood sugar should receive IV insulin infusions to maintain the serum glucose less than 180 milligrams per deciliter for the duration of their ICU care."

Again, they gave it a Class 1A recommendation.

So, after discussions with the Technical Panel, what we were trying to figure out was a way to capture some information about blood sugar control in the ICU without too much burden for the hospitals. So, the recommendation of the Technical Panel, along with representatives of the Society of Thoracic Surgeons, is that we simply ask the hospital to capture a data element that looks at whether any of the blood sugars between 18 and 24 hours were greater than 180 milligrams

per deciliter.

So, we would only look at a single timeframe. We would not look at two days, as we have in the past.

The feeling of the Committee, and particularly the input of STS, is that by 18 to 24 hours after surgery the blood sugars should be controlled, and the Class 1 recommendation is that those blood sugars be below 180.

And so, I know Tori Lynch and Wanda Johnson are on the call. They can perhaps give you a little better description of the data elements that we are testing. But it would simply look at whether any blood sugars in that timeframe, 18 to 24 hours after surgery end time, were greater than 180 milligrams per deciliter.

Wanda, do you want to add any comments?

MS. JOHNSON: The only thing that we also added is, if there was no blood sugar

documented between 18 and 24 hours after anesthesia end time, that they would record the highest one between 12 and 18 after, just in case there were no blood sugars at the 18-to-24-hour mark.

MR. BRATZLER: So, that is our recommendation for the changes. That met the approval of our Technical Expert Panel and, again, had a great deal of input from the Society of Thoracic Surgeons, the guideline panel that gave the Class 1A recommendation for blood sugar control.

CO-CHAIR MORRIS: Does the group feel as though our conditions are satisfied?

MEMBER CIMA: I just want to clarify a couple of things. First of all, that was not Class 1A evidence they cited. They cited, "All patients with diabetes undergoing cardiac surgical procedures should receive an insulin infusion in the operating room and for at least 24 hours postoperatively to maintain serum glucose levels less than

180." That is evidence Level B. So, just to make sure we are clear on what they said.

MR. BRATZLER: But look at Recommendation No. 7.

MEMBER CIMA: I am looking at the paper right now. The recommendation is not a Class 1.

MR. BRATZLER: I read it. I read it.

MEMBER CIMA: I have it right here in front of me.

MR. BRATZLER: Yes, well, Recommendation No. 7, I read it verbatim. "Patients with and without diabetes with persistently-elevated blood glucose should receive IV insulin infusions to maintain serum glucose less than 180 milligrams per deciliter for the duration of their ICU care. Level of evidence A."

MEMBER CIMA: Well, if you look at the second page of that recommendation, which I mean I am reading it right now also, Level

4, management of hyperglycemia using insulin protocols in the perioperative period, recommendations Class 1B.

So, "All patients with diabetes undergoing cardiac surgical procedures should receive an insulin infusion in the operating room for at least 24 hours postoperatively to maintain serum glucose levels less than 180 milligrams per deciliter (level of evidence equals B)."

MR. BRATZLER: But, yes, so I see what you are reading, but that, again, looks at the perioperative period which includes the operating room, which is not what we look at. We only look at the postoperative care in the ICU. So, Recommendation No. 7 is where we based our --

MEMBER CIMA: It says for at least 24 hours postoperatively. I just want to clarify that. It says that right there, postoperatively for 24 hours.

MR. BRATZLER: Well, exactly, but

so does Level 7, or Recommendation No. 7 is during the ICU stay. So, that is all we look at. We do not look at interoperative control.

MEMBER CIMA: Okay. So, the only issue is, what does it mean? Why are they going from 200 to 180 now?

MR. BRATZLER: Basically, because we felt like we should align with the published guideline. We, frankly, always try to do that. We always try to make sure that we incorporate guideline recommendations for our performance metrics.

MEMBER CIMA: Now what if the patient -- there are exclusions here for patients on inotropes. Is that going to be an exclusion built into it or are they just saying any patient?

MR. BRATZLER: No, it is all patients. And you're absolutely correct. We recognize that inotropes drive the blood sugar up, but that is all the more reason that those patients need to have IV insulin infusions to

keep their blood sugars controlled.

As was pointed out by the STS Committee, the inflection point for mortality, morbidity, and infections goes up once the sugars are in excess of 180, whether it is induced by inotropes or whether they have diabetes. So, it is for all patients. We don't exclude patients on inotropes.

MEMBER CIMA: The only thing I am a little confused about is that you are saying that patients -- it says in item 7.1, it says, "Patients with and without diabetes with persistently-elevated serum glucose should receive an IV insulin infusion." So, that's Level A evidence.

But the question is, so persistently-elevated, does that mean that you are saying that all ICU patients need to have an insulin infusion to make sure or is it patients that are persistently elevated? I mean, because the question here is you are saying you want to introduce a measure where

you have to put everyone on an insulin infusion to make sure they are on insulin for 24 hours or you are saying they have had a blood sugar greater than 200, which may not correspond to persistently elevated.

The question is here, is exactly what are you talking about; what do you want us to do? So, are you saying that everyone has to be on insulin infusion or everyone has to be on it even when the evidence says it is persistently elevated, not a single evaluation point?

MR. BRATZLER: Right, right. Well, the recommendation for the Committee was that the blood sugars between 180 -- I mean, sorry -- between 18 and 24 hours should be 180 or less. And so, we don't require anybody to be on insulin infusion. I will make that every clear. We don't look at how patients are treated. We allow the hospital to control the blood sugar any way. And if the patient's blood sugars are less than 180, there's no

need for any intervention at all.

But the recommendation of the Committee was simply that the hospital record the highest blood sugar between 18 and 24 hours. And we looked to see whether or not it was greater than 180 milligrams per deciliter.

MEMBER CIMA: So, if they check it one time and it is 200, that is the first time, they are going to get penalized, and, then, they start the insulin drop appropriately? I mean I am just concerned about the language. It is persistently elevated. You are looking at a one-time measure.

MR. BRATZLER: Right, I understand your point completely. I think it would be a rare case that somebody would check for the first time a blood sugar 18 hours after cardiac surgery, but --

MEMBER CIMA: No, but, then, if it is --

MR. BRATZLER: -- I understand

your issue.

MEMBER DUTTON: Arguably, if the first time they check it is 18 hours after surgery, they should be dinged because that would be out of compliance with --

MEMBER CIMA: No, but what if is fine multiple times and, then, one time it is not? I mean, which one are you going to take? Say they have two of them, and one is higher and the next one within 24 hours is lower because they have adjusted the insulin.

MEMBER DUTTON: It seems to me that there is going to be a tradeoff here between the burden of data collection and the ability to assess the measure, to figure out if we are actually doing the right thing.

MEMBER CIMA: That's the point. Is it a good way of measuring it? Should we say evidence of attempting to maintain insulin? I am just saying it is a very difficult measure to meet by many hospitals that are keeping them on insulin drips. And

also, the way it is written is difficult to manage.

You know, the data suggests not a single blood sugar level. The data, all the data suggest persistent elevation over a period of time. This is not a good measure for measuring that.

MEMBER KLEINPELL: Well, you know, last time we had a lot of discussion how, as it was raised before, with a 6:00 a.m. focus, that a lot of hospitals were, in fact, putting patients on intensive insulin therapy in order to meet that 6:00 a.m. measure.

So, now we are moving to 18 to 24 hours, which, again, is more comprehensive, you know, definitely better than just the 6:00 a.m. But I think we also have to look at unintended consequences.

You know, we talked before about some people potentially gaming the system to try to get that 6:00 a.m. within normal values, and, then, maybe not adhering to

looking at the value and how that varies over time.

I think the change is definitely, again, more comprehensive, but I guess feasibility may have to be considered. I guess we won't know until we really see this in practice.

Is there any data with respect to longitudinal monitoring of glucose that you have that would relate that many patients are within the range by 18 to 24 hours?

MR. BRATZLER: I personally don't have, we don't have any of that data from the SCIP project. I believe that Tony Furnary's group and others do have considerable data on that particular topic. I would have to defer to them.

You know, I think one of the issues that we struggle with here is that we all recognize, as with the current measure, with the 6:00 a.m. measure as it exists currently, occasional cases will fall out even

with appropriate care. We understand that.

We have never, never ever said that the target of performance on the measure is 100 percent. So, we haven't tried to build the performance measures to address every possible rare clinical exclusion to the performance measure.

So, I think it is possible that a single blood sugar between 18 and 24 hours could pop up above 180 and a case might fail. I guess that is a possibility. And we recognize that.

We know from cardiac surgery data that 50 to 60 percent of all cardiac surgery patients are hyperglycemic postoperatively. So, it is a big population of the cardiac surgery group that should receive insulin therapy, and it has to do with the prevalence of diabetes and the use of inotropes, and all the other things that happen to these very large operations that are occurring to these patients.

So, that was the recommendation of the Committee with a great deal of input from the Society of Thoracic Surgeons' Committee, also.

You know, we talked about it. I know this group has talked about it. We have talked about it before.

Ideally, we would capture every single blood sugar and calculate a daily average, but that simply isn't feasible.

MEMBER DUTTON: So, Dr. Cima, you are concern is that a single value doesn't accurately reflect that integral over time? Or that the rate at which it doesn't reflect it is high enough to invalidate the measure?

MEMBER CIMA: Well, I just think their own literature says that it needs to be a trend of persistently elevation. And that is GSS, despite how you may want to parse the words. But, certainly, it is better than two days at two separate times at 6:00 a.m. It is a better measure, but I just think it is not

really directing at what the literature and the scientific evidence support, which is better glucose control over a period of time.

I think in the ICU setting, as they physicians who care for these patients, you can have a patient on an insulin drip and have them at 150 for six hours straight and check another blood sugar and it is 200 for no reason. I think that is what people feel, that they are really tightly controlling these people.

And there is evidence in the literature, Class 1 evidence, that says extremely tight control in ICU patients and cardiac surgical patients increases their risk for death and sepsis. So, it works both ways.

And the data says a trend and persistent elevation. We are looking for a single marker in critically-ill patients who are on inotropes who have had major interventions, who are ill patients coming in. The vast majority have diabetes, as was

mentioned. I think it is not the ideal measure for it.

Now does that invalidate it? I think it is better than it was, but I still think it doesn't meet the scientific evidence that is available now, that it is a persistent elevation of blood glucose levels greater than 180 that have been shown to increase the risk of infection and mortality. It is not a single value.

CO-CHAIR MORRIS: Okay. So, our role right now is to determine whether or not our condition was satisfied.

Dale, would you please repeat the numerator statement for the group?

MR. BRATZLER: So, the numerator would be patients having cardiac surgery whose blood sugars -- so, what we do is we capture the highest blood sugar between 18 hours and 24 hours. So, it would be those patients who had a blood sugar that was 180 or less in the 18-to-24-hour timeframe, recorded as their

highest sugar.

MEMBER HALPERN: So, are you saying that somebody is looking at all the sugars anyway?

MR. BRATZLER: They are looking at those between 18 and 24 hours. They only capture one. They just look at the chart for that timeframe.

MEMBER HALPERN: So, why aren't they, then, looking at the average, if they are looking at them all anyway?

MR. BRATZLER: Well, that would require them to capture them all. We don't do that. So, this just lets them glance at the lab or the glucometer records and capture a single value.

MEMBER HALPERN: Okay. Well, then, I go back to Dr. Cima's point. If they look at it and they see the majority of the glucose, all of them except one is high, is that really a good measure? Because you do have, if you force people to overtreat and

they become hypoglycemic, again, there's consequences to that.

MR. BRATZLER: I completely understand that. But, again, I want to make it very clear. We have never, never ever promoted intensive insulin therapy, never, with this performance measure. As we have now or as it was in the past, we have never promoted it.

The clinical endocrinologists published guidelines about a year ago focused on all ICU patients, not just cardiac surgery. That recommends for all ICU patients blood sugars be maintained between 140 and 180. That is the national recommendation for ICU patients. The STS guideline is consistent with that, that the blood sugar simply be maintained less than 180.

So, believe me, I understand the comments about unintended consequences, and I have made it very clear in multiple forms that we do not promote intensive therapy. So, what

we are trying to do is figure out a way to capture reasonable control of blood sugar in a timeframe after surgery when most, the expert panel, the STS felt that by 18 hours the blood sugars should be reasonably controlled.

MEMBER CIMA: I have a question, then, hypothetical. If at 18 hours the blood sugar is 150, well, but let's say, for clinical reasons, they do. And, then, at 18 and a half hours it is 200. And, then, at 19 hours it is 150 because they corrected it. Do we fail them?

MR. BRATZLER: Yes, the case would fail as recommended.

MEMBER CIMA: So, is that the clinical, is that what you want? You want us to be treating that as opposed to the clinical circumstances? Is that what you are asking?

MR. BRATZLER: So, I understand the point, and that is why I made the point. We understand that sometimes patients who

receive appropriate care fail the measure. We understand that. The target is not 100 percent. It is high, and right now the national average around the current measure is in the 90 percent range, but some cases still have blood sugars that are elevated.

So, again, I am just reflecting the conversation of the Committee to try to balance this combination of data burden with having a measure that is reasonable and consistent with the guideline.

CO-CHAIR MORRIS: Dr. Martin?

MEMBER MORTON: I was going to say that I think the proposal is a step forward. It is something that is different from what was encountered before. I think all of us reflected some of the frustration we see in clinical practice in having to meet an arbitrary 6:00 a.m. time. So, I think the proposal is really a step forward in that it is capturing it within a more reasonable timeframe.

I think the other point is that we all agree that blood sugars should be controlled after surgery. I don't think anybody has a quibble about that.

I think the one point here is whether or not a single time point is reflective of the entire post-op glycemic state of the patient. I think that is really the question. I guess I think that is one that is a little hard to figure out. There is always a balance of trying to get as much data as possible as well as having the parsimony to get things actually done.

But I think the proposal is a step forward, where we don't have that arbitrary 6:00 a.m. time.

MR. BRATZLER: So, I guess I am going to ask the question, what would be enough data? Should we say two or more blood sugars greater than 180?

I mean, what we are trying to avoid is extensive data collection, data

points. We have used metrics like this for data abstraction in the past where the abstractor can scan a record, and if they see a number that is above a threshold, they stop looking. They don't have to look anymore.

So, once we go to two or three data points, then we start to require them to look at every single sugar that occurred during that timeframe.

MEMBER CIMA: Could you do something like 18 to 24 hours, if there is a blood sugar greater than 200, was there a correction and another blood sugar within four hours that was less than the goal? I mean, could you look at it that way? So, it is a two-tiered system.

I agree with John; I think this is a much better measure in the sense it is more realistic in the sense of what we are trying to achieve. But could you say, within 18 to 24 hours, if it is less than 180, fine, we're done. If it is above 180, was there an

intervention to correct it in a reasonable time period and bring it down below 180?

MR. BRATZLER: Well, I am going to defer to you or Tori to think about the mechanism of capturing that. It would require additional data elements for sure.

MS. LYNCH: Right, and as we look forward to EHRs, how would you write that? If you found a level higher than 180, you would then look for an entry for a correction? And you would then look for an entry for a controlled value.

To me, it is a matter of hospital data abstraction burden. But, yes, we can write specifications to address that.

MR. BRATZLER: We can certainly explore it.

MEMBER SIPERSTEIN: I would just like to echo the comments that, yes, this measure I think is moving in the right direction. One of my concerns is that these measures are not just passive observations of

a best clinical practice, but the institutions, quote, "study for the test". So, it does influence practice pattern.

And you could very easily see some practice patterns evolve that are no good. So, if you get one blood sugar that is 179 at 18 hours, are you going to quit measuring it for the next six hours? I mean institutions are going to do things to game the system, and I think it is important, as Bob has hinted, to craft the measure so that it really is encouraging a best clinical practice, you know, somewhat analogous to the normothermia measure, in that after surgery that if you may not meet the temperature requirements, at least you have tried.

In this way, I agree that some type of two-tiered measure where you do go just above the threshold, and you have demonstrated success/attempt to correct, then that would mean that the institution is moving towards or doing their best to institute a

best practice.

CO-CHAIR TORCHIANA: I just had a thought pursuing this same line, because the idea of just checking one sugar and then not checking another until 24 hours and 30 minutes sounds too, it is a little too cynical to support a practice like that.

But what if there were a sugar that was greater than the threshold, whether that is 180 or 200, and that is the trigger? And, then, if there are other sugars prior to 24 hours that are below the threshold, it basically gets erased? So, the trigger creates a tension, and only if the subsequent sugars up to 24 hours remain elevated does the measure actually become a failed glucose control. And you don't have to look for the intervention. The intervention would be implicit in the lower sugar.

MR. BRATZLER: I think that would be easier to do.

CO-CHAIR MORRIS: I think what we

are doing here is adding additional conditions, which is absolutely fine, but Melinda please correct me if I am wrong. So, that is something that we are sort of talking about as a group, putting additional conditions on this before we would vote on it.

MS. MURPHY: And that is fine, but it will make a difference. We will not be able to vote on it today because, again, I am going to make the assumption that they will need to go back to their technical panel to have this discussion. So, we are talking about it moving out some period of time.

CO-CHAIR MORRIS: Okay.

MR. BRATZLER: For just a specification change, we can figure out a way to communicate via email with the panel. So, we can expedite the discussion.

CO-CHAIR MORRIS: Okay. I think that sounds fine.

Anything else that anybody wants to say about whether the conditions were

satisfied on Measure 0300 before we move on to voting on the other measures?

(No response.)

All right, let's go to page 2 of the agenda. The next measure is 0114, risk-adjusted postoperative renal failure.

We had a lot of discussion about this previously. Does the group wish to have this represented or do you folks believe that you are ready to go ahead and vote?

(No response.)

Okay, let's go ahead and vote.

Okay. So, the first question is, do you recommend this measure for endorsement, yes or no?

(Vote.)

Are we all set?

That is 17 votes for yes, two for no, and three abstain. So, we do recommend it for endorsement.

Sorry. One for no, one for abstain. Yes, I definitely need eyes in the

back of my head more than ever before.

Measure No. 0115, risk-adjusted surgical re-exploration, is there a desire among the group to redescribe this measure or to represent this measure?

MEMBER CIMA: Did we talk about this one last time? I forgot. I don't remember it.

CO-CHAIR MORRIS: I am blanking, my memory is blanking out on this one as well. Did we actually talk about it?

MEMBER FINDLAY: So, we punted on all these. Why, again? Could you remind us why?

MS. MURPHY: The items that are here that you are looking at right now, each had some condition or request for additional information from the Committee to the developer. And, then, on the 31st of March --

MEMBER FINDLAY: Right.

MS. MURPHY: -- those conditions were discussed in terms of whether or not they

met the request requirement of the Committee.

And the ones that you are looking at on the 31st all met the conditions that were requested. So, now the last piece, the last step, is to vote on whether or not you would now, having that condition satisfied, are ready to vote on a recommendation for endorsement.

MEMBER DUTTON: I believe most of the STS ones, the issue was disparities, right? Presenting disparity data, which they did on the phone.

MS. MURPHY: And what you see on the screen up here in the shaded box is what the condition was, how the developer responded, and what the action of the Committee was. So, in each case, what you are going to see projected is what your condition was, the response, and your determination.

(Vote.)

CO-CHAIR MORRIS: Okay. So, the voting is completed. There are 19 votes for

yes, none for no, and one abstain.

The next measure is Measure 0116, anti-platelet medication at discharge.

As Melinda pointed out, the conditions will come up in the shaded box.

There are a number of conditions, and during our conference call we agreed that they were met. I think it is time to go ahead and vote. Does anybody want to discuss further?

(No response.)

Okay.

(Vote.)

There are 20 votes for yes, none for no, and none for abstain.

Measure No. 0118, anti-lipid treatment at discharge.

Again, this is around disparities. Anybody want to discuss it?

(No response.)

Okay.

(Vote.)

Can everybody hit their button one more time on their remote?

(Vote.)

Twenty votes for yes.

Measure No. 0119, risk-adjusted operative mortality for coronary artery bypass graft.

And again, you can see in the shaded box the conditions.

Is everybody ready to vote?

(No response.)

Okay.

(Vote.)

Nineteen votes for yes, one vote for no.

The next measure is Measure 0120, risk-adjusted operative mortality for aortic valve replacement.

Again, we are looking for information on disparities, and we will go ahead and vote.

(Vote.)

Nineteen votes for yes, one vote for no.

Measure No. 0121, risk-adjusted operative mortality for mitral valve replacement.

Again, questions about disparities, and we will go ahead and vote.

(Vote.)

Okay, could everybody hit the button on their remote one more time?

(Vote.)

Nineteen votes for yes, one vote for no.

Item No. 0122 is risk-adjusted operative mortality for mitral valve replacement and coronary artery bypass graft surgery, and we will go ahead and vote.

(Vote.)

Let's go ahead and hit the buttons one more time. This is sounding familiar.

(Vote.)

Okay. Nineteen votes for yes, one

vote for no.

Measure No. 0123, risk-adjusted operative mortality for aortic valve replacement and coronary artery bypass graft. Okay, we will go ahead and vote again.

(Vote.)

And that's 20 votes for yes.

Measure No. 0129 was risk-adjusted prolonged intubation, ventilation.

And our conditions were consideration of a change in the time limit to less than 24 hours and data on disparities. We, in the conference call, believed that our conditions were met.

Does anybody want to discuss this?

MEMBER ROGERS: Yes. This is Terry. I would just like to make one comment.

I am still opposed to 24 hours. Notwithstanding the complexity of many patients in surgery, I think that 24 hours gives some leeway and latitude. So that, if you actually had a sloppy method of addressing

people postoperatively, it gives you too big a window, in fact, to practice that sloppiness. So, I would just register that comment.

CO-CHAIR MORRIS: Thank you.

Anybody else?

(No response.)

Okay. Let's go ahead and vote then.

(Vote.)

It says 15 votes for yes, four votes for no, and one abstain.

One of the issues that arose in our last discussion that we kept coming back to was the issue of topped-out measures, measures in which the criteria had essentially been met to an extensive degree, maybe not 100 percent, but pretty close to it. And this was one of those measures.

We will need to, basically, determine how we want to deal with that because several of these measures may be

topped-out or approaching being topped-out. I guess if they are approaching it doesn't count, but if they are topped-out, then we need to make a decision about whether we want to continue to endorse them in the maintenance measures.

Is there more?

MS. MURPHY: So, the discussion at the time of the last in-person meeting was that performance was at or above 90 percent, but there was discussion that the Committee felt that compliance should be closer to 100 percent.

So, the remaining question is, do you feel that, based on the information that you have here, that it should remain in an active status until it gets closer to 100 percent?

MEMBER HALPERN: You are talking about the intubation one still?

MS. MURPHY: Right.

MEMBER HALPERN: The intubation

one that we were just talking about a second ago. Because you have the sterna wound infection up on there, I think it is confusing people.

MS. MURPHY: Yes, this is where the discussion was. So, you're right, this had the voting for the other one, but what you need to be looking at is the information that reflects the discussion.

MEMBER HALPERN: Certainly, if they keep it at 24 hours, I think we should get closer to 100 percent.

MS. MURPHY: And that is not at issue. I mean it is as specified. So, that would remain.

MEMBER FINDLAY: So, at the last meeting we had, did we, in fact, move any measures? I don't recall moving any measures to inactive status. We did? How many?

MS. MURPHY: There were not any that were recommended to move to inactive status. There was discussion of a few of them

that could potentially be candidates for inactive status. But, really, the inactive status -- and, actually, I am going to let Heidi speak to this, if she is willing to do so -- is a new approach that NQF is taking, and is still working that through.

But the draft guidance out at this point is that you might determine that a measure could be moved to an inactive status only on the condition that it is important to measure, it meets all of the criteria, but the performance is so high that you would not expect that there would be a huge opportunity cost to increase it beyond the performance level at the time.

MS. BOSSLEY: Right. So, I think what you will want to look at is not only is it topped-out at the top, you want to look at the variation. You also want to look at the disparities data, if they are able to provide it, because if you see there is an issue with the disparities, you may very well determine

that you don't want to use the inactive status here.

The one caveat I will say is this status is going to the NQF Board for their consideration because it does need to be approved since it is a new endorsement status. That will happen next week.

So, if you do decide that you want to use this and it is appropriate for any of these measures, we will caveat it with "if the Board approves it". And, then, if they do not, then we will bring it back to you. And, then, we will have to figure out what your next step will be.

But, again, we don't want to remove endorsement from measures that may be very appropriate to continue to surveil, to look at over perhaps a periodic time period. So, this was our way of acknowledging that these are still evidence-based measures. They meet all the criteria except they don't show any disparities, there is no variation, that

type of thing.

CO-CHAIR MORRIS: So, just to really clarify, measure that are really good measures that we are concerned providers potentially may not attend to any further if they come off of the active list. If they are still in inactive, they would be easily reinstituted.

MEMBER FINDLAY: So, at the deliberations next week, is NQF going to create a process for the Committees, this one and others, to move measures to inactive status, which might include a number of steps that you would take to do that?

MS. BOSSLEY: Right. So, right now, what we are thinking is anytime measures that come up for maintenance, because it would only be applicable to --

MEMBER FINDLAY: Right.

MS. BOSSLEY: -- measures that are currently endorsed, it will be brought to the Committee as a potential option.

The other thing that we are building in is the ability to continue to capture this updated specifications on those measures. Because if we do think it is important for periodic surveillance, that type of thing, we will continue to have that on our website and available for individuals.

So, that is our preliminary thinking at the moment. Whether that would mean more in the way of guidance to the Steering Committees, we are developing that now. And again, your discussion will help us with that. So, we appreciate it.

MEMBER FINDLAY: So, there is a broader context here that there is a lot of pressure at NQF and also at CMS and in other venues to remove topped-out measures. That happened in the hospital value-based purchasing final rules in the last week or two, where there was pressure put on CMS to remove four or five of the measures, and they did so.

So, I guess I would say that we ought to be paying more attention to that and build it into the process. I guess today and tomorrow, if we have the option, if there are measures that look to be topped-out, then we have the option of sort of punting on them until there's clarification. Correct?

MS. MURPHY: I am not sure what you meant by punting. You have the opportunity --

MEMBER FINDLAY: Well, put it aside not to vote, not to take a vote at this time.

MS. MURPHY: Or what you have the opportunity to do is go ahead and vote --

MEMBER FINDLAY: And, then, ask for --

MS. MURPHY: Right. But you would vote, you would first make the determination --

MEMBER FINDLAY: Right.

MS. MURPHY: -- that it remained

important and it met all of the criteria. So, you would go through all of that process and, then, because of the level of performance, which is the only kind of measure that is deemed to be important, and it is going to have the opportunity to move to an inactive status.

MEMBER FINDLAY: Right.

MS. MURPHY: You would recommend that occur contingent upon approval by the NQF Board for that designation.

MEMBER FINDLAY: Thanks.

CO-CHAIR MORRIS: So, the question for us with this measure, we voted on it, and that part is done. But the question is, do we want to recommend that it move to inactive status? Do we think that it is topped-out?

MEMBER ROGERS: Well, one of the reasons that it might top out is that the threshold is too generous. I mean, hopefully, one learns from these. This is what we are all about. What we are trying to do is change

behavior. And if we have a threshold that is so generous that nobody's behavior is going to change, then putting it on inactive status is not doing justice to the reason we have it to begin with.

So, I would think that, harkening back to I think 24 hours is way too generous, and it is set up as sort of a surrogate for other issues. For instance, if more of your patients require or if enough of your patients require more than 24 hours, should maybe they have been operated on to begin with? I mean that is one question.

And secondly, to keep it at this because it might lead to premature extubation makes no sense to me at all. I mean you extubate when you need to extubate.

But the way it is set up, I don't think we are learning anything from this. Just putting it on inactive status doesn't get us where we need to be. It is sort of shuffling it under the rug. We need to have

a threshold that actually delineates or discriminates between, and helps people look at their behavior. It makes them change. And 24 hours for me just doesn't cut it.

CO-CHAIR TORCHIANA: If I could comment on this one, we talked about this on the call. It is an interesting measure. I would say that the reasons why patients are intubated for more than 24 hours after heart surgery are multi-factorial.

The one that is most obvious is to get at the anesthetic technique, and an anesthetic technique that results in prolonged somnolence and delayed extubation is probably not a good practice.

But patients also sometimes stay intubated because of hemodynamic instability7 or hemorrhage. The thinking, at least on the call that I was on, was that this might be augmented by an average time to extubation for non-outlier patients, which would be a good measure of the anesthetic approach and the

approach to the care of the patient. And, then, the greater than 24 hours intubation would become more of a surrogate measure for the issues like hemodynamic instability or operating on the wrong patient, somebody who has already been on a ventilator for a month when they go to the operating room, for example. And having those two together might complement and solve some of the problems that we are concerned about with this measure.

CO-CHAIR MORRIS: As a group, we voted to endorse the measure.

I hear what you are saying, Terry, about that the fact that you would like for the measure to be a better measure. In terms of whether we believe that it should -- but we voted to endorse it as a group. Do we think that it is premature, then, to put it on inactive status?

MEMBER ROGERS: Yes, I really like David's approach. That way, at least it discriminates. You learn something about the

group who maybe should have been intubated earlier, and you learn something about the other group who, say, had complications. To lump them all together I think doesn't do service to either one of those groups.

So, maybe a message we can give back to STS is to help us maybe craft another, if that is something we can recommend, to address that very issue that David brought up.

MS. MURPHY: So, then, the recommendation is that this remain in an active status, that it be endorsed, and that STS get a recommendation that they develop a companion measure. Okay.

CO-CHAIR MORRIS: That sounds good.

Let's move on to -- oh, is there more?

MR. SHAHIAN: Dave Shahian from STS. I came in late because I had travel problems this morning.

Could I just ask for some

clarification as to what you are looking for? I understand you are asking for a change in the measure, which I thought had been endorsed. So, could I just ask for some specifics?

CO-CHAIR MORRIS: We are asking for development of a companion measure.

Go ahead.

CO-CHAIR TORCHIANA: Yes, I think, David, not a change in the measure, but a companion measure that might get at the question of anesthetic in postoperative practice, such as median time to extubation for those not intubated for greater than 24 hours, something in that mode.

Because there is a concern that this measure was described as a sloppy measure that allows for a lax practice, that allows for patients who ought to routinely be extubated before 24 hours remaining intubated for a longer period of time than is appropriate.

MR. SHAHIAN: So, you are asking for a new additional measure?

MS. MURPHY: Yes.

MR. SHAHIAN: Thank you.

MEMBER DILLON: When does this measure come up for review again? In three years? So, would it be a fair statement that the likelihood of this being passed three years from now, certainly in its current state, is probably highly unlikely, given the concerns in that?

Because what I am trying to set up is a timeline or an expectation for the development of a new measure.

MS. MURPHY: So, your additional recommendation is that this companion measure be brought forward at the next time this is brought forward for maintenance review?

MEMBER DILLON: Yes.

CO-CHAIR MORRIS: Okay. The next measure, 0130, risk-adjusted deep sterna wound infection rate.

This condition was also around data regarding disparities. Let's go ahead and vote.

(Vote.)

Nineteen votes for yes, one abstain.

Measure No. 0131, risk-adjusted stroke/cerebrovascular accident.

And our conditions were data regarding disparities; also, the time period in which the cases were eligible for inclusion in the numerator, and exclusion of patients with a prior stroke. And STS responded, and we considered that to be sufficient.

Any other discussion of this?

(No response.)

Let's go ahead and vote.

(Vote.)

Could everybody hit their button one more time?

(Vote.)

That's 20 votes for yes.

The next measure is Measure 0300 that we have spent some time talking about. We decided that the conditions were not entirely satisfied, is that correct? Not entirely satisfied; we asked for more conditions.

Let's move on to Measure 1501, risk-adjusted operative mortality for mitral valve repair.

So, our conditions were that we wanted clarification whether this is simply open-chest procedures, and the response was that it applies to all mitral valve repairs. And, then, we asked for data on disparities.

Any discussion?

(No response.)

Okay, let's vote.

(Vote.)

Please hit your buttons one more time.

If there is somebody that consistently notices that, after they hit

their button, the vote is over, please come up and see us. Maybe there is a battery change.

(Vote.)

Twenty-one votes for yes, and no votes for no or abstain.

Do we have 21? Do we have 21? Okay.

The next measure is 1502, risk-adjusted operative mortality for mitral valve repair and coronary artery bypass graft surgery.

Again, we wanted further data on disparities. Let's go ahead and vote.

(Vote.)

Twenty-one votes for yes, none for no or abstain.

So, we are done with that portion. Now it is time for us to move on to discussion of Phase-I-related and competing measures and the Steering Committee recommendation for endorsement. It is, yes, time to move on to that discussion.

Was Alexis going to talk about this? Were you going to talk about --

MS. MURPHY: So, this activity is a new piece of work with NQF. And it is looking at, again, where we have a number of measures that are related and the extent to which those measures can be harmonized, assuming that they both are important or whatever number are important.

And when you are looking at related measures, you are looking at they may have the same measure focus or they may have a different measure focus but the same population or care setting. So, they are not looking at the same focus and the same population. The relationship does not put them in competition one with the other.

So, the first question to answer, while you are looking at the measures, is, are they, in fact, truly, as you see them, related measures that should be looked at, then, further in terms of harmonization, where it is

appropriate to harmonize, where there are unintended differences that could, in fact, lend themselves to harmonization?

And, then, you are looking at whether or not the measures are competing. Are they, in fact, looking at the same measure focus and the same population, such that you want to or we should be looking at whether or not we should have two measures that are looking at the same focus in the same population, and whether or not one of those measures is a best measure. This goes to the best-in-class question.

So, what we have got are a set of measures that you will look at in terms of here what you see on the screen is the measure related to a database for cardiac surgery. This is Measure 0113 that you have evaluated in Phase I.

And a related measure is another STS measure looking at participation in a database for general thoracic surgery.

And, then, the third measure is the one that Helen mentioned to the group when the group met in March in terms of a generic database.

So, those are the ones that you are now looking at in terms of related or competing measures.

The threshold question that the group discussed at its in-person meeting was the question -- and I think this was also discussed during the conference call -- was the question about clearly and certainly at the time of initial endorsement of measures related to participation in database. It was at a time whenever there were few measures, when there needed to be a mechanism through which information could be collected in order to learn about performance.

And one of the issues that the group discussed is whether or not that time had gone by and whether or not the information that now provides data for measures that are

drawn from a database supplant the need for a measure that says participation in a database.

So, there are two questions, and Dr. Cima is not happy?

(Laughter.)

MEMBER CIMA: No. I am trying to clear that up.

MS. MURPHY: He's concentrating? Okay. So, you can address it either way. You can take the discussion of should there be measures around participation in a database, any database? And if the answer to that is yes, then it is probably good to go on to the next part of taking a look at these measures in terms of harmonization and competing measures.

CO-CHAIR MORRIS: One of the things that we are thinking about here that has come up in multiple conversations in the group was the precedent this will set because there are more and more databases being developed, regional databases, national

databases. And there is more interest in looking at data in secondary databases for quality improvement and accountability.

So, we are not making a final decision here, but we are making some important preliminary decisions that likely will be discussed and revisited many times in the future.

And there are number of different ways to handle this. I think that the first question, as Melinda said, the first question that we should probably discuss is whether we think that participation in any sort of systematic database should be essentially mandated, so any sort of database should be mandated.

Anybody want to open that up?

MEMBER HALPERN: Well, I will tell you that your Board certification in surgery now depends on you having some sort of database, you know, some sort of quality control of yourself. So, it essentially

forces you already to participate in a database.

CO-CHAIR MORRIS: Do we mandate that hospitals participate in databases?

MEMBER ZAMBRICKI: I was trying to remember the measures for cardiothoracic surgery. Are there any that do not include a requirement for comparing a performance against a database? Because that was one of the issues you brought up. Is this already included in the other individual measures? Do you see what I am saying?

CO-CHAIR MORRIS: Yes, I think that they are all within the STS database. Is that right?

Okay. All of the measures were within the STS database. And so, one of the questions that arose was, is it appropriate for us to mandate that hospitals participate in the STS database?

MEMBER ZAMBRICKI: Yes, and I thought we had decided, no, that it had to be

some type of a database, multi-hospital. So, I thought that that was the change for those individual measures, that it could not dictate STS database.

MS. MURPHY: But I think what we got was clarification that the measures, while it is clear that they say the information for those measures that is presented in terms of performance was drawn from data in the STS database, that participation in the STS database was not a requirement of the measure.

MEMBER ZAMBRICKI: Right.

MS. MURPHY: And there was discussion around what might you then have to do in order to put that data together. But in none of those measures, in the measures themselves, none of those require participation in the STS database.

MEMBER ZAMBRICKI: Do they require participation in a database?

MS. MURPHY: And we are talking now about the measures using data derived from

the database. We are not talking about these measures.

MEMBER ZAMBRICKI: Correct.

MS. MURPHY: And so, say your question.

MEMBER ZAMBRICKI: Well, I am just wondering, is it duplicative to have a measure that just calls for participation in a database if the individual measures for performance include in the measurement of that measure the necessity to be participating in a database? So, by nature of the fact that you are submitting information on those measures, you must be already participating in a database. Does that make sense? In which case, we wouldn't need any of these because anyone submitting a cardiothoracic measure would be participating in a database, if I am understanding it correctly.

MS. MURPHY: And that is the threshold question. The issue, though, is that they would not necessarily be

participating --

MEMBER ZAMBRICKI: Right.

MS. MURPHY: -- in the STS database.

MEMBER ZAMBRICKI: Right.

MS. MURPHY: But they would have to be able to derive, to collect and have all of the information for computing the performance against the measure.

MEMBER HALPERN: I think also saying you are participating in the database makes you have to compare yourself to others, as opposed to just creating your own database. So, I think it is a separate issue.

MEMBER MORTON: That is the same point I was going to make. There is a real utility in participating in a national registry where you can benchmark against other people, rather than just having a home-grown registry.

MEMBER DILLON: Yes, but I agree with Christine that, if we are to focus just

on the particular measures, by that happening people are going to have to participate in a database. They will have to be comparing themselves because our measures will require that, particularly as they evolve in their complexity.

So, I am not sure that we have to say, you know, it is mandatory that you have to participate in a database. We are going to say you have to submit these measures. How you get these measures and how you wish to compare yourselves to others, as required by the measure, is up to, should be up to the institution. Because, otherwise, how many databases -- you know, where does it stop? How many databases do we tell people that they apply to?

MEMBER HALPERN: But are the measures actually saying that you have to compare yourself? I don't think they are actually saying that. They are saying that you need to collect this data, but it is not

really saying that you have to compare yourself.

MEMBER MORTON: I would kind of err on the side of being explicit rather than implicit in the registry. I know what you are saying, Peter, but it is possible people could interpret it in a way where they don't participate in a national registry. So, I would rather be explicit about it.

MEMBER FINDLAY: What is the burden of that, this measure, on providers, on collection of the data, on anything? It a yes-or-no question, essentially.

CO-CHAIR TORCHIANA: Well, there is a history here that Dr. Shahian may want to comment on as well. So, the STS database is overwhelmingly the most ubiquitous benchmark database in the country for cardiac surgery. But there are two very important other databases that are highly respected, the Northern New England database, which was a spontaneous effort among New England

hospitals, and, then, the mandatory New York State reporting system, both of which had long lives and have been very effective. And they use all the same concepts. The comorbidities and the outcomes overlap, but are non-identical.

And certainly, everyone involved in the STS database would acknowledge that institutions in New York State that are in the New York State database, and from a mandatory requirement by New York State government, or those that are voluntarily in the NNE, are doing a good job of doing a benchmark quality assurance.

And so, that is why this language I think ought to be ambiguous, that it is in a dataset where you are benchmarking against other institutions.

David, do you have anything to add to that?

MR. SHAHIAN: No, I think you have covered it correctly. We are now at about 95

percent participation nationally. The only programs that don't participate are basically those in northern New England or New York.

When the measures were initially adopted, we were well below that level, and this was an attempt, using a structural measure of quality, to drive participation in a national registry.

For all practical purposes, given the pressure to report the hard outcomes that we now report, you are correct that de facto you must belong in a systematic cardiac surgery registry or you basically can't exist in cardiac surgery in the United States today.

Now, having said that, that is not true for the general thoracic, where I would say a large percentage of general thoracic surgery in this country is conducted in programs where there is no participation in a systematic database.

So, the measure for cardiac surgery, do what you feel is correct. I don't

think it is going to change the status quo because we are at a point in cardiac surgery right now where you can't survive without being in a registry.

CO-CHAIR TORCHIANA: If I could just make one other comment on this, the requirement presupposes the existence of an effective organization to do the data analysis and benchmarking. And requiring institutions to participate in an organization that doesn't do that effectively is probably not a wise thing to do.

CO-CHAIR MORRIS: That sort of gets to your question about what is the cost to the hospital or provider.

MEMBER MORTON: I had a question maybe for the STS representative about, isn't there a requirement to give reimbursement to participating in STS at this point for a lot of the payers?

MR. SHAHIAN: Many payers do actually require this, either for

participation in their plan, certainly for premium status of various sorts. So, yes, that is correct.

And did somebody want to ask about the cost of participation? Was that a question? Yes?

MEMBER FINDLAY: Yes, the burden and cost.

MR. SHAHIAN: Certainly. The average cost to a participant, and a participant is typically either a hospital or a large surgical group, the cost is approximately $3,000 a year. On top of that, you have software costs which range anywhere from about $10,000 for a very, very good program anywhere into the hundreds of thousands for all the bells and whistles. But you can do it for about $10,000 software cost. The major cost, as in all registries, is the cost of data collection. We estimate one FTE for every 500 cases entered.

MEMBER CIMA: I mean that is a

major issue that was brought up even at my institution. It is getting to, like Peter was saying, how many do you do?

We have four full-time abstracters for the STS database at the Rochester campus. We have abstracters in Florida and Arizona. Basically, the decision at Florida was, well, we were also paying for an abstracter for NSQIP. The multi-specialty NSQIP came out. So, we are going to stop participating in the STS. We are going to use the cardiac surgery NSQIP.

So, there is this dynamic, as we have mentioned multiple times, the burden of data abstraction and participation in different databases. I mean the cost to actually submit the data is relatively mild in medical terms, $10,000 maybe. But abstracters, trained abstracters, a full-time FTE can run into the multiple tens of thousands, if not, with benefits and everything, $100,000 a year. So, you have to

get some real value out of it.

If they can do 15 cardiac cases for the NSQIP sample as opposed to doing 40 STS ones, because they have a requirement to do all, I believe, in STS, and I am not 100 percent, people are making their cost/benefit thing. It is a huge number. As it becomes more and more burdensome, it is going to become more and more of an issue.

MEMBER DUTTON: To get back to the fundamental question, I think there is value to participating in a registry above and beyond simply counting your cardiac mortality or anything else. And for less-involved specialties than cardiac surgery, they are at a much more primitive point that David mentioned of having very low participation in national or aggregate benchmarking.

CO-CHAIR MORRIS: This sort of gets back to what Dr. Halpern was saying. That is that participating in a registry means that you are comparing yourself to other

hospitals, and that is kind of the bottom line; whereas, the other measures that we voted on just mean that you are measuring those items.

So, do we believe that we want to advocate for essentially endorsing that hospitals belong to a registry where they are comparing themselves to other hospitals?

MEMBER SIPERSTEIN: Yes, I think we have had the discussion before. I think everyone is on the same page in that registration makes sense. I think most hospitals realize that, yes, there are going to be resources that need to be put towards this end.

The difficulty has to do with multiple similar, quote, "competing" sources. And obviously, if you are participating in multiple registries, you are diluting resources that could be better spent in different areas.

CO-CHAIR MORRIS: Okay. Sorry.

So, the decision here I think is, are these measures competing? Are they related? If they are related, should they be harmonized? So, competing versus related is a decision that we have to make here.

MEMBER ROGERS: I'm sorry, I just have one other comment or question. If we move towards a mandate, which it sounds like we are inching towards, that normally carries with it some sense of authority or jurisdiction, i.e., punishment if you don't belong. I am not sure that we actually own that kind of --

MS. MURPHY: And I don't think that that is what it is trying to do, and that is not what this question is trying to go to.

MEMBER ROGERS: Okay.

MS. MURPHY: The threshold question is, does endorsement continue for a measure that speaks to participation in a database?

MEMBER ROGERS: Okay. Because I

heard the "mandate" word two or three times. I just wanted to -- right.

CO-CHAIR MORRIS: I think that this question is trying to go there, but we don't necessarily have to help it.

(Laughter.)

MEMBER ROGERS: Okay.

MEMBER CIMA: A question that arises, though, from this is, what I remember the discussion was with STS last time was, yes, it comes from the STS database, the data to support that there is a performance gap and whatever, and that there is also these are our measures.

But, then, the question became, and then they said, well, it is all on the web. People can go to our website and download the way of doing it so that you get the number, the risk-adjusted number. It is basically saying, if we endorse a measure designed that way, yes, you don't have to participate in STS, but you have to follow

STS's rules in how they adjust their risk adjustment for whatever it is.

That is where I fundamentally had the problem last time. Because what if there is a different risk adjustment model used by a different nationally-qualified clinical database registry? Should we say you can participate in that one, but for this specific measure you have to use the rules and the methodology outlined by the STS that we have endorsed.

Even though you participate in this measure, you participate in a nationally-qualified clinical dataset that comes up with, that evaluates cardiac surgery, but just doesn't happen to use the same methodology, you still have to use that methodology. That is where you get into the problem.

If you are going to say we need to harmonize, then you have to harmonize the measures and say that the methodology for the measurement is identical or standard, so that

you don't have to use the STS methodology. Where risk adjustment is very different, very important in cardiac surgery, they have a different risk adjustment than maybe NSQIP has.

So, we are saying, yes, you can go to the web and do it, but what that means is I've got a person in my hospital that has got to do the NSQIP dataset and, then, I have to take our raw data, I have to go to the STS dataset. I have to get their methodology. I have to redo it, then submit it to you.

That is a data burden, and that is what we have to clarify here. Are we saying you participate in a dataset and, then, the measures could come from STS, but they can't use the STS risk adjustment because that means, in essence, you have to do what they are doing. That is picking a winner.

MEMBER SAIGAL: I agree with that point, although there is some benefit to homogenous reporting of data across all these

datasets. These are the best minds in the field in NSQIP and the STS. If there are differences in their models, maybe an appropriate thing is to have them come with their measures and debate that, why one is better than the other, when they are putting their measures forward.

MEMBER CIMA: So, then, you are saying, either way, let's say at the end of that debate A wins versus B. Then, you have to say we are letting you participate in database. We say that is fine for that measure. But, at the end of the day, that data really doesn't count; you have to use somebody else's. You have to use their methodology.

What's the purpose of doing that then? Then, you are saying: well, then, why are you paying for this? You should just do that.

MEMBER SAIGAL: Well, I am saying you can agree with one of the developers that

their model is stronger, and I think that it is an appropriate thing for us to do, to have data nationally be collected in a similar way and be comparable.

MEMBER HALPERN: I think, if I get you right, what you are trying to say is, if one database actually has a better model, then all databases should move toward that model.

MEMBER CIMA: I think that is a different discussion, though, than measurement development. We are being asked to develop a mortality, you know, risk-adjusted mortality for aortic valve replacement, AVR 0120. Although it based on STS risk-adjustment methodology, so if you use your NSQIP to report your risk-adjusted operative mortality for mitral valve, is that sufficient? Or do you have to say I collected that data, I have done it, and now I have to go back and use the free, publicly-accessible data methodology on the STS website to reprocess my data and re-administer and then resubmit it? That is what

I am saying, there is a difference.

CO-CHAIR MORRIS: Okay. Dr. Cima, thank you.

I think that I agree that that is sort of a nuance, the discussion about whether other databases should move toward one best model, and we talked about those nuances as well previously, but that is sort of outside of what we are trying to accomplish for right now.

Dr. Shahian, did you have something to add?

MR. SHAHIAN: Yes, just a point of fact. I have served as an advisor to NSQIP for the last four years. I serve on their Measurement Committee, their Policy Advisory Committee.

There is no NSQIP cardiac surgery model. There is no competing model. I work very closely with Cliff Ko, who directs it. We work in a complementary fashion. They have no desire to get into cardiac surgery. They

do not have a cardiac surgery module. They do not plan a cardiac surgery model.

And feel free to call them anytime to check on that.

MEMBER CIMA: Well, Allan and I sit on the Joint Commission NSQIP Board, and in the multi-specialty there's cardiac surgery in it.

MR. SHAHIAN: There is also prostatectomy and laminectomy and neurosurgery. These are not models designs for cardiac surgery. These are generic risk models. They were never intended to be used for something as specialized as this.

And they are actually phasing out that model as they bring in specialties. Specific models in general surgery, but they have no plans to do so for cardiac surgery.

MEMBER CIMA: But, right now, they are reporting on that. You can participate in a multi-specialty and have cardiac surgery patients included in that.

MR. SHAHIAN: Well, you can include anybody you want, but the risk models are so generic they would be, in our opinion, of little value.

MEMBER HALPERN: That actually comes up on vascular because I have noticed that in NSQIP predicted mortality of certain cases, like, for instance, an infected aortic graft, their predicted mortality of the patient that was involved was 2 percent, which is ridiculous.

MEMBER CIMA: Well, I am not supporting one or the other, but I am just saying, as Dr. Shahian just pointed out, in his opinion and the STS's opinion, it is not any good.

But we are asking whether or not participating in a national database, a systematic database, is appropriate. That is the question at hand here, not whether or not one is better than the other.

So, there is a national

clinically-recognized database that does have cardiac surgery data in it. Whether you agree with the modeling or not is a different question.

My point is this: should a hospital, then, also need to participate in STS or can they use the cardiac, fulfill the requirement for cardiac surgical patients if they participate in NSQIP?

MR. SHAHIAN: The question begs credulity. Nobody is using NSQIP for cardiac surgery. It was not designed for cardiac surgery.

You may choose to use it, but you would be alone, essentially, in the country. I can't say anything else. It is a matter of fact.

MEMBER MORTON: Just as kind of a historical point, NSQIP went through a high-risk model just to look at certain cases. That is being phased out, but they are going to come up with surgery-specific models. I

agree, right now, I guess there is not a plan to do something around cardiac, but I don't think it precludes that they will.

But, in just reading the measure and talking here, it seems fairly agnostic. It doesn't say that it has to be STS. Am I reading that correctly?

So, I think the main point all of us want to have here is that we do believe it ought to be collected. We believe it ought to be benchmarked. It ought to be risk-adjusted.

Does it have to be STS? I think that is the only question that is out there because STS seems to have the preponderance of the hospitals, but there can be other datasets that come up.

So, I think, from my sense, I am fine with it as long as it is not exclusionary to just STS.

MEMBER DILLON: But doesn't it become focused, aren't we, then, directing it, if we approve this, because of the

stewardship?

MEMBER MORTON: I don't know. That is a good question. I don't know.

MEMBER CARPENTER: Isn't the question at hand whether these are competing measures? I mean these have all been approved. We are not asking whether to approve these or not. Whether these are competing, and in that competing, make us have to select one versus the other. I would suggest that these are not competing, that these are related, that they are different patient populations. Obviously, the measure on the right is a much more generic measure for a hospital, and every hospital can meet that in a variety of ways, and that these are not, as listed there, competing, and really have nothing to do with a lot of the discussion that has been which database is actually used.

And so, I would suggest that, unless someone thinks that they are competing,

and we should discuss that, that we call these related and move on with that.

CO-CHAIR MORRIS: Does everybody agree to call these related rather than competing? Dr. Cima, do you agree?

So, another issue that we have here, then, is that since 0113 is a maintenance measure, and Dr. Shahian said that there's more than 95 percent participation, and that if you don't participate, essentially, you're alone, should this become an inactive measure?

MEMBER DUTTON: Yes, I think I suggested that last time. I think in many, many specialties and disciplines this is a very important measure. I mean the number of anesthesia practices reporting anesthesia data is less than 5 percent right now. So, there is a huge opportunity in other specialties, but in this case they have topped it out.

CO-CHAIR MORRIS: Anybody else?

MEMBER ZAMBRICKI: I think that

that is a good point because isn't part of the intention that these measures will be publicly reported and the public can make judgments about quality based on the result of a measure? So, in a simple way, a lot of people would get credit for quality in the public's eye, when it may actually be the status quo.

CO-CHAIR MORRIS: Any other thoughts on 0113 and whether we would recommend that it go to an inactive status?

MEMBER WILHOIT: Another perspective on this would be that the rate is 95 percent, which means that 5 percent aren't participating, and do you want to draw attention to that 5 percent? Is that an issue?

Because it is not the rate that is reported; it is the individual facility or practice. Does this practice, does this hospital participate in the registry, not that the rate is 95 percent.

MEMBER STAFFORD: Do we know what

the number is combined for all of the databases that are out there? So, if it is 95 percent for STS, of that 5 percent gap, how much of that are those that belong to the New England Consortium and the New York State databases?

CO-CHAIR MORRIS: With the language of the measure, they would still be adherent to the measure.

MEMBER ZAMBRICKI: But it sounded like, from the STS representative, the compliance is about 100 percent if you include the New England database and the New York, from what he was saying earlier.

CO-CHAIR MORRIS: Does anybody disagree with recommending that this move to inactive status?

(No response.)

Okay.

MS. MURPHY: So, then, what we can do is vote for retaining endorsement with the caveat that it would be on inactive status if

the Board of Directors approves that, the status.

MEMBER MORTON: Can I ask what that means practically? Is it like double-secret probation when you go to inactive?

(Laughter.)

What does that mean?

MS. MURPHY: It means that it would be set with measures that would not be expected to be monitored and reported on -- because, remember, this is a public reporting piece -- on the same frequency, but that it could be brought back into an active status and, then, reported upon without having to go through all of the process.

MEMBER HALPERN: When you say not with the same frequency, does that mean they still get monitored at some frequency?

MS. MURPHY: One, it is voluntary and there is the expectation that there would be some monitoring of it, but periodically -- there would not be an expectation on the part

of NQF to have, unless the Board of Directors specifies something, to have a specific interval during which it is reported upon.

MS. BOSSLEY: Right. This is where it gets challenging because we don't control the implementation. We are not necessarily involved in the implementation and the reporting of these measures.

So, anything that does go into inactive status, in the same as anything that is endorsed, other than us saying this meets our criteria, that is an influencer on who implements and uses these measures. But, beyond that, we can't say with all certainty that any measure that is endorsed will be used and reported.

So, inactive, in that way, I think that is something that the Committee needs to balance as well. When you do have a measure that moves into inactive you are sending a message, still valid, still reliable, still important, topped-out, that may actually move

a measure off of implementation programs. I think that is part of what you need to look at, is what are the unintended consequences of that. With this measure, I don't think that is an issue. You may have a few coming forward that you would need to address that.

MEMBER KLEINPELL: Well, I have a question. When a measure is moved to inactive status, is there some rationale then provided, so that when people go there, they see the reason that we are putting it inactive is because there is 95 percent to near 100, if you include all of them? So that they know we are not saying this is not important anymore.

MS. BOSSLEY: So, this is part of the messaging that we are still figuring out because we are new to the maintenance piece of this. We want to be able to message, when removal of endorsement occurs, what was that reason. We want to be able to say why it moved into inactive status.

So, we are currently figuring out,

what is that? What is the level of detail? But, again, inactive would clearly state it meets all the criteria except it is topped-out, yes.

MEMBER STAFFORD: So, if you are still working on the language, should we even be talking about this at this point? Because you are now asking us to discuss and vote on something that may change, and we have just spent a whole lot of time talking about this.

And so, if we don't know what the message is going to be based on our vote, then it really doesn't seem like the right thing to do at this time.

MS. MURPHY: Can I respond?

MS. BOSSLEY: Sure.

MS. MURPHY: We provided information about inactive status in terms of the draft. And that is one of the reasons why you would caveat a decision. But within the information about inactive status, it provides for the fact that the measure focus continues

to be important, that it meets the criteria based on evidence of scientific acceptability, usability, feasibility, and that the only issue is that it is topped-out.

So, that part of the criteria I don't think is at issue or would be expected to be changed by the Board.

MS. BOSSLEY: Right. I mean we have put the inactive status proposal out for comment and received comment from the membership. In general, it is supportive. I don't anticipate that the Board will have an issue with it.

So, in that way, we feel that it is probably in the best interest of your time to at least consider it now, rather than have you come back on a conference call and do it. But we are taking a bit of a chance asking you to do this. We are also doing it with another committee as well.

With the language, though, on the website, let me be clear, all your documents

that go out for comment and in the report will clearly explain what you have proposed and what your recommendation is. Nothing else on the website will change.

The measures stay endorsed until this final review is done. It goes through the Consensus Standards Approval Committee, and, then, it goes to the Board.

That gives us until November or so to figure out the language that actually goes on the website with these measures that are inactive. So, I think we are okay. I mean, from a staff perspective, we feel comfortable. And so, the question is, if you all feel comfortable having that conversation today, but we don't think that there will be an issue going to the Board.

MEMBER MORTON: Can I ask a question? I am just curious what Dr. Shahian has to say about it. Will this have an impact on anything going on in cardiac surgery if we go to making this measure inactive?

MR. SHAHIAN: No, I don't think this will have any impact because the measure is derived from database participation or, essentially, required if you want to have a cardiac surgery program.

The only thing I would worry about was the comment that was made, what is the message we are sending? I am a little concerned about that. But I leave that to your good judgment.

MEMBER DUTTON: Success, that is the message. You've done it. You've got everybody.

MR. SHAHIAN: Well, as long as everybody understands, as long as the rest of the world understands that, I think it is all in how it is presented.

CO-CHAIR MORRIS: Maybe the name could be changed from inactive to successful.

(Laughter.)

MS. BOSSLEY: We have had I can't say how many high achievement, emeritus. We

don't know the right name.

(Laughter.)

Any suggestions are welcome.

MEMBER COLLINS: I have a question. On these inactive states and for measures that are topped-out, is there a concern with regressing, regression? High rates of compliance, 95-plus, it is not such a big deal here, but I see measures we will talk about in the future where I think it is, where if you go back to 90, 80 percent, you know, some even lower rates, that is a real concern.

Have there been similar scenarios where things have been moved to inactive, the Committee has determined it is topped-out, and, then, you have seen a regression?

MEMBER AFSAR-MANESH: Well, I know this is not really up for discussion at this point, but it would be very interesting, as this gets developed by the staff, that perhaps there is an infrequent reporting component

built in. And again, maybe for this one it doesn't need to be as frequent, but for some of the other ones, again, that are coming up today, something that looks at it on an annual basis, or whatever frequency seems reasonable, so that it is inactive, but still there is some feedback loop. So that, if it slipping, we would know about that.

MS. BOSSLEY: They are all really good points, those things that we have been trying to figure out because no one knows what the potential impact of this is. So, this is, again, where we need to balance because we do not control who implements/reports out all of this. That is not within NQF's purview.

But, as we move toward electronic systems, we are hoping that this actually can be much easier. Especially with some of these measures, I think once you have an electronic system, you will be able to poll this and be able to surveil it periodically. But it is one of those things we honestly don't know the

answer to yet, and we will have to continue to monitor with it, work with CMS, HHS, and others, to figure out what we can do to assist.

CO-CHAIR MORRIS: Okay. So, we essentially agreed as a group that there is an inactive or emeritus or some other label for Measure 0113, we agree that it should go there.

We should move on to Measure 0134.

MS. MURPHY: We should vote.

CO-CHAIR MORRIS: Do we need to officially vote? Oh, okay, we are going to officially vote.

Do you recommend this measure for endorsement?

MS. MURPHY: And the caveat here is it will also say "and moved to inactive status".

MEMBER DUTTON: So, we are recommending it be moved to inactive status, what this really should say?

MS. MURPHY: As an endorsed measure.

CO-CHAIR MORRIS: Endorsement does not preclude inactive. In fact, it has to be endorsed to move to inactive.

(Vote.)

Please hit your button again.

(Vote.)

It is 20 votes for yes and one abstain.

The next measure is 0134, surgery patients who received appropriate venous thromboembolism prophylaxis within 24 hours prior to surgery to 24 hours after surgery end time.

And again, we are going to see measures that are potentially related versus competing here.

Oh, I'm sorry, did I misread that? I'm sorry. Yes, I am getting ahead of myself. I apologize.

Coronary artery bypass graft using

internal mammary artery, 0134.

MEMBER DUTTON: Why does STS have two of these measures? What is the difference?

CO-CHAIR MORRIS: That is what we are supposed to be talking about here.

MR. SHAHIAN: I think the difference is simply, is mainly the exclusions, and 0134, actually, what you probably have in your paperwork, our staff discovered a slight error, which I hope will be corrected.

But the appropriate exclusions list should not include the IMA is not a suitable conduit. When this was discussed recently on a call, we agreed to eliminate that exclusion. So, the exclusions that are considered valid are previous cardiothoracic surgery and released on a radiation emergent or a salvage procedure or the absence of LAD disease.

So, again, it is mainly the

exclusions. I think 0134 is the measure that is being put forward.

MEMBER WILHOIT: The endorsed Measure 0516 was endorsed four years ago? Wouldn't that be up for review again by now if it is four years old? I was confused by the date on that.

MS. HAN: This is Jane Han from STS. May I answer that question?

Measure 0516, actually, it was first introduced in 2008 and it received time-limited endorsement. And so, Heidi Bossley is there, and she can probably speak from an NQF perspective.

But it was reviewed by the CSAC last July, July 2010, and received final endorsement or full endorsement as a result of that discussion.

The reason why, as Dr. Shahian explained, the exclusions are listed differently, which we will have corrected, but the other difference is that 0516 is at the

physician-level while 0134 is at the hospital- or facility-level. I don't think we have any problems combining them into a single one.

The reason why we have a physician-level measure is for PQRI. So, I don't what happened in 2008. It was before I started here. But I think it was submitted as a separate measure with the same specifications just for that reason.

Does that make sense?

CO-CHAIR MORRIS: Yes. Thank you.

Would the group like to recommend that they be harmonized?

(Chorus of yeses.)

Okay. Anything else?

Dr. Shahian, do you want to say anything?

MR. SHAHIAN: No.

MS. MURPHY: So, we would not vote on the measure as it exist today because you have just recommended that it be harmonized. So, you would be voting on something that is

going to be overtaken very quickly.

CO-CHAIR MORRIS: Okay. The next measure, 0218. Now we are going to talk about surgery patients who received appropriate venous thromboembolism prophylaxis within 24 hours prior to surgery to 24 hours after surgery end time.

MEMBER CARPENTER: I can make a couple of comments because I have looked at this.

We have been through one of these measures, the 0218, last time, and we voted to endorse that. That is specific to selected surgical patients, and there is a list of diagnoses that apply in their measure.

The other measure, the previously-endorsed measure from the Joint Commission, is medical and surgical patients, essentially all patients over the age of 18, as I read it. So, it is a much broader inclusion criteria. The 0218 is a subset of the patients included in 0371. So, that is one difference.

The other difference is what is considered VTE prophylaxis. In my understanding, the 0371 measure includes any documented VTE prophylaxis measure or technique, medication or mechanical prophylaxis, versus the 0218, which requires really the chest position guideline measures primarily to meet the criteria.

So, the 0371 is not restricted to any one individual set of guidelines. There are competing guidelines, chest positions, which is based on DVT prophylaxis. There is one from the orthopedic surgery group which is based on symptomatic PE, which differ from the chest position guidelines.

So, those are some of the differences. It is a smaller, it is a subset of the bigger group, and it is a more specified degree of what is acceptable prophylaxis.

There is some controversy over what is acceptable prophylaxis, and those two

groups have not been able to work out mutually-standardized prophylaxis guidelines. So, there are sort of competing guidelines at this point.

MEMBER HALPERN: Correct me if I am wrong, but I think some of the controversy arises from bleeding versus clotting risks, the risk of bleeding into recently surgical spaces and the detrimental effects of that.

MR. BRATZLER: Yes, this is Dale Bratzler.

I need to correct something here.

MEMBER HALPERN: I can't hear you.

MR. BRATZLER: I was the Chair of the Technical Expert Panel for the NQF/Joint Commission VTE measures that are being discussed here, the competing measure.

The surgical population is explicitly excluded from that measure. That measure focuses only on medical issues. So, this measure focuses on surgical patients, based on the HACCP guidelines, which gives

procedure-specific Level 1 recommendations for VTE prophylaxis.

The populations do not overlap with the Joint Commission measure. We worked carefully to make sure that they did not overlap.

And the reason that the NQF-endorsed Joint Commission VTE measure allows any forms of prophylaxis is because for medical patients there are almost no guidelines that are explicit about which forms of VTE prophylaxis are appropriate for medical patients.

So, I chaired that Technical Expert Panel and know the Joint Commission measures very well. And I can tell you without question that the SCIP population is excluded from those measures.

MEMBER STAFFORD: Dale, we are sitting here looking at Measure 0371, which is the Joint Commission measure, and it includes surgical patients, both in the numerator as

well as the denominator. It says, "Medical and surgical inpatient discharges. If surgery and incision time is greater than 24 hours of admission, patients must have documentation of prophylaxis within 24 hours of hospital admission."

That is the measure that was endorsed in May of 2008.

MR. BRATZLER: And I am telling you that, if you look at the denominator specifications, the SCIP population is excluded.

MEMBER STAFFORD: No.

MEMBER CIMA: Well, in our document it says, "Denominator, all patients".

MEMBER STAFFORD: That is correct.

MR. BRATZLER: Well, I can tell you that, if you look at the ICD-9 codes, the SCIP population, we worked carefully to make sure that the two measures did not completely overlap. So, they don't overlap. The SCIP population is not included in the Joint

Commission measures.

MEMBER DUTTON: Yes, on page 22, it does list SCIP as an exclusion.

MR. BRATZLER: And the reason is that there is a large number of surgical patients that don't fall under the SCIP population for which there aren't explicit recommendations around VTE prophylaxis or others. They are hospitalized and they have some surgery.

But the SCIP population, the SCIP measure that you are discussing now only reflects those operations for which there are published Level 1A recommendations for VTE prophylaxis. And that is the big difference between the Joint Commission and the CMS measure. The Joint Commission measure looks essentially at hospitalized patients. Did they get prophylaxis or did they have documentation of a reason to not get prophylaxis, whether they are medical or surgical? But the SCIP population is excluded

from the Joint Commission measure.

MEMBER STAFFORD: Perhaps even more of a point to harmonize these, because I would say that most hospitals and most surgeons are doing this for all of their patients, whether they are in the SCIP population or not.

MR. BRATZLER: And I understand that point. Perhaps there is a way to do it, but the difference here is that we look at the appropriateness of prophylaxis based on surgical procedure in the SCIP population; whereas, the Joint Commission measure only looks at whether any form of prophylaxis was given at all.

You know, there just aren't many people out there that would recommend TED hose for hip replacement surgery or other things. That is the big difference in the specifications of the measures.

The Joint Commission measure does not look at appropriateness, and that is the

big difference for the SCIP measure. It looks at appropriateness.

CO-CHAIR MORRIS: So, Measure 0218 is really for the SCIP population, it sounds like; whereas, the other measure specifically excludes the SCIP population.

Jim, I want to hear what you have to say. So, I am not trying to push us ahead, but we will need to decide whether we want to endorse 0218 and, then, request harmonization from the other measure.

MEMBER CARPENTER: What I was going to say, I think I understand that better with Dale's clarification. Because I don't think we have all the details in what we were presented, and the Joint Commission makes it particularly hard to follow their algorithms. If you ever look at those, you just give up pretty quickly. And some of the tables are a little hard to find.

So, I think that clarification is helpful, and perhaps they are not the same

group of patients. They are really more related.

I guess having different criteria for different groups of patients for data collection/extraction does complicate that process. And if there is an opportunity for harmonization, that would be an advantage.

You do lose potentially some of the appropriateness. As we discussed before, we did endorse that, but there is controversy and guidelines do change with time. So, there may be an opportunity for harmonization here.

MEMBER SIPERSTEIN: I just want to comment, having had the pleasure of just re-reviewing the chest guidelines, it is a very thick tome. Part of the reason is it is not a simple algorithm. And subset of patients in whom there are data to make more specific recommendations exist, but there are also large groups of patients where we really are working in the realm of generalities.

So, I can understand the concept

that, where the data exists to be more prescriptive about types of VTE prophylaxis, even the chest guidelines, there are menus given for each of these categories, not a single treatment. It may make sense, where there is data, yes, to be more prescriptive, and where there is not data, to simply be more general in terms of recommending some type in the absence of more definitive data.

MR. BRATZLER: Yes, so I think you have highlighted why there are two separate measures. And again, I chaired the panel fore the Joint Commission. That is really the population.

We are focused on the hospitalized population, medical patients, surgical patients that don't fall under the SCIP population. A hospital can pass with any form of prophylaxis or documentation of a reason not to use prophylaxis.

The SCIP measures are very procedure-specific. In the development of the

denominator, we went through every ICD-9 code, now ICD-10 code, to define which patients would fall into the population of the denominator for each type of surgery. And, then, the menu, the menu that you mentioned from the chest guidelines, is the appropriate menu that is used to put the patient in the denominator for each type of surgery.

CO-CHAIR MORRIS: Okay. Thank you.

Are we ready to vote on whether to endorse 0218? Anybody want to say anything else about it?

(No response.)

All right. Let's go ahead and vote. Be sure you aim at Jessica.

(Vote.)

Is there somebody missing?

Please press your buttons one more time.

(Vote.)

Okay. Seventeen votes for yes,

two votes for no, one abstain.

Our next decision is whether we want to request that this be harmonized, whether we want to request that the other measure be harmonized with this measure.

Are we actually voting on that or just discussion it?

MS. MURPHY: It is not a voting --

CO-CHAIR MORRIS: It is not a vote? Okay.

MS. MURPHY: But it would be a recommendation from the --

MEMBER RIVENBURGH: It sounds like that is no, since they are defined by different populations. And the reason they are defined as different populations is that the standards are different or the guidelines are different.

CO-CHAIR MORRIS: In that case, what I would like to request is just that this be clarified a little bit more when it comes back up for review. Are the populations

actually different?

Jim, did you have anything to add?

MEMBER CARPENTER: No.

CO-CHAIR MORRIS: So, the next measure is Measure 0360, esophageal resection mortality rate, risk-adjusted.

And this was to be evaluated with 0361, right? I think we are talking about whether to harmonize.

MS. MURPHY: Yes. So, this is looking at 0360 and 0361, mortality and volume, in comparison to the Leapfrog measure, which combines the two into a single measure.

And the question here is not whether it competes, but, rather, is there an opportunity to take 0360 and 0361 and in some way, through combining them or otherwise harmonizing with the Leapfrog measure, and it was endorsed just in last September.

CO-CHAIR MORRIS: One of the issues that came up in our discussion of this was that volume is more predictive of

mortality than mortality is because the volume can be so low in some centers. So, that should be part of our discussion here.

MS. MURPHY: Go ahead and have that discussion, whatever they want to do.

CO-CHAIR MORRIS: All right. I guess what we should really discuss here is, do we want for these to be harmonized before we actually vote on them?

MEMBER HALPERN: Yes.

CO-CHAIR MORRIS: Anybody disagree with that?

MEMBER WILHOIT: I think there's reasons not to harmonize. I mean for the point you just raised of volume being an important measure. And if you look at the Leapfrog measure, as best as I can figure out from the details that are here, it is all a calculation.

And so, you have no direct measurement. You are not quite sure what you are getting. All you get is the bottom line.

And if mortality is less predictive than volume, I like seeing the numbers rather than just seeing a calculation.

CO-CHAIR MORRIS: Anybody else want to respond to that?

(No response.)

I think that mortality has a lot of street credibility. And therefore, it is important in that way. It probably should be measured. If it is measured as a completely separate entity, that can be deceptive.

Is it possible to have a volume measure and, then, to have a mortality measure that is harmonized with volume as a separate -- can we make that recommendation?

MS. MURPHY: Yes.

CO-CHAIR MORRIS: What do you guys think about making that recommendation?

MEMBER WILHOIT: And the recommendation there being to harmonize 0360 and 0361, but keep them separate from the Leapfrog measure? I think that has a lot of

credibility, to do that.

MS. BOSSLEY: This is Heidi.

I mean one option is to recommend that they be paired, which means that they are endorsed as two separate measures, but should be implemented and used together, which is, I think, what you are trying, where you are headed with those two.

MS. MURPHY: And when they were initially endorsed, they were endorsed to be reported as pairs.

MS. BOSSLEY: So, I think you just want to continue that endorsement.

CO-CHAIR MORRIS: Any other discussion of this? Carol, do you feel satisfied by that?

MEMBER WILHOIT: That would sound good to me.

CO-CHAIR MORRIS: Okay. All right, let's go ahead and vote.

(Vote.)

Please aim your hand-helds at

Jessica and vote one more time.

(Vote.)

Okay. That is 20 votes for yes, no noes, and no abstains.

I think what we are going to do now is, it's 11:19, yes, we will take a break. Maybe is five minutes too short? A 10-minute break? And, then, we will come back.

Yes?

MEMBER CARPENTER: Let's finish this because they are a pair.

CO-CHAIR MORRIS: Oh, I'm sorry. Oh, I beg your pardon. I'm sorry about that. This is what having no eyes in the back of my head leads to.

On 0361, let's go ahead and vote. Please just go ahead and vote twice.

(Laughter.)

(Vote.)

That's 20 votes for yes, no noes, and no abstains.

And now a 10-minute break,

assuming nothing else pops up behind me. So, we will see you back here at 11:30.

(Whereupon, the foregoing matter went off the record at 11:20 a.m. and resumed at 12:40 a.m.)

CO-CHAIR TORCHIANA: Could we get underway again, please?

Okay. Before we get to consideration of candidate measures, Melinda has a comment to make on the prior discussion.

MS. MURPHY: This is just finishing up the discussion that you just had with respect to the esophageal resection measures.

What we will do, what we need to do is be able to get you the full specifications, probably in the form of the measure evaluation for the Leapfrog measure, so that you can look at that in terms of best measure evaluation against the two that you just recommended for endorsement. And we will get that out to you, and we will have that

discussion on a conference call, when we talk about the other related and competing measures.

CO-CHAIR TORCHIANA: So, the consideration of candidate measures will begin with a presentation by the measure developers.

And I would ask that we try to be quite concise in this, so that we can actually get to some of the measures before our lunch break at 12:40.

And so, I am going to follow the order that is on the agenda and begin with the Society of Thoracic Surgeons. I think we have eight or nine developers. So, we will start with the STS.

David, are you going to be the presenter?

MR. SHAHIAN: Yes, and I can be quite concise.

This is a measure designed to promote the use of beta blockade in the preoperative cardiac surgery patient.

Cardiac surgery entails some manipulations of the heart, fluid shifts, electrolyte shifts, use of cardioplegic/cardiopulmonary bypass, and a number of other unique features that result in a particularly high incidence of postoperative atrial fibrillation which can be quite a morbid complication and certainly a costly complication.

The use of preoperative beta blockade primarily for this purpose has been a longstanding ACCAH Class 1 indication. There are now more than 60 randomized trials, excuse me, 30 randomized trials that show on average a 60 percent reduction in perioperative atrial fibrillation with the use of preoperative beta blockade.

There are other reasons to consider beta blockade in the cardiac surgery patient, including one study by Bruce Ferguson based on several hundred thousand patients in the STS database showing a slight reduction in

mortality in patients with ejection fractions over 30 percent, substantiated by two smaller observational studies. There are 10 randomized trials showing a reduction in the incidence of postoperative VT and DF.

This is a measure where we have considerable room remaining. There is a gap and there is variability in performance. We have in our more recent iterations recognized the fact that there are patients for whom preoperative beta blockade may not be possible, for reasons of hemodynamic instability, for example. And we now allow such exclusions if they are documented in the medical chart.

An audit of this measure is included as part of a routine audit, and our auditors have found that there is agreement between the use of contraindications and what is available in the chart to document those reasons in over 95 percent of cases. So, that seems to be working.

Now there may be some concern due to the fact that the recommendations have changed in general surgery. There has been a move away from routine preoperative beta blockade.

A little bit different situation. First of all, most of the studies that have shown concern have been studies in which fixed dose agents have been used a very short time before the procedure and without titration, which is clearly not the appropriate way to do it. Very salutary results have been found where patients have been titrated to an optimal heart rate, and the agents have been started well in advance.

So, I don't think that the concerns in --

CO-CHAIR TORCHIANA: David, if I could ask you to be just a little more high-level?

MR. SHAHIAN: I'm done.

CO-CHAIR TORCHIANA: We have a

bunch of these to get through.

Any comments on the beta blocker at discharge that seems pretty self-explanatory?

MR. SHAHIAN: This is preoperative, I think we are doing.

CO-CHAIR TORCHIANA: No, I am saying there is a second STS measure.

MR. SHAHIAN: Oh, yes.

CO-CHAIR TORCHIANA: Any comment on that one?

MR. SHAHIAN: The same reasoning, except to say that virtually all, by definition, cardiac surgery patients would qualify for one of the ACC/AHA recommendations for secondary prevention for use in patients that have MI, unstable angina, heart failure.

So, there are compelling reasons, not the least of which is long-term survival in a number of studies, to support this measure.

CO-CHAIR TORCHIANA: Thanks very

much.

The next measure is from CMS, and I believe Dale on the telephone is going to be presenting as the developer.

MR. BRATZLER: Yes, this is Dale. Let me pull up my notes.

This is the conversation about the beta blocker measure? I want to make sure I have got the right one.

CO-CHAIR TORCHIANA: Yes, the beta blocker measure, surgical patients on beta blocker prior to admission.

MR. BRATZLER: Right. So, the denominator for this measure is restricted to patients undergoing surgery who are on beta blockers prior to arrival. So, the American College of Cardiology, the American Heart Association have, it is a Class 1A recommendation continuing beta blockers in patients who take beta blockers at home.

So, the denominator population for this performance measure is patients having

surgery who take a beta blocker at home. We recently recommended some mild changes to the performance specifications to improve the measure and to reduce some of the data burden.

So, the change to the measure now is that the post-op length of stay is greater than or equal to two days. And we look to see if the patient received a beta blocker on the day prior to surgery or the day of surgery. So, no longer are we requesting a 24-hour timeframe. We just simply say, did the patient get the beta blocker either the day before or day of surgery, and, also, did they get a beta blocker on either post-op day 1 or post-op day 2?

One of the problems with the measure in the past is that the hospitals could pass the performance metric by simply giving a single dose before surgery. And now, you know, the real focus of the guideline is to continue beta blockers in patients who take them at home. So, now we will look to see,

did they get a dose the day of the day before surgery, and if they are in the hospital for greater than two days, did they get dose on either post-op day 1 or day 2?

CO-CHAIR TORCHIANA: Thank you. Could you also say a word on the measures on the next page of the agenda, hair removal, complication rates following arthroplasty, and readmission rates?

MR. BRATZLER: Yes. So, I cannot speak to either the arthroplasty or the readmission rates. We don't have responsibility for that measure.

The hair removal measure, I think this has been discussed before. This is a measure that I think there has been some discussion about approaching very high rates of performance. The performance measure basically looks at whether the patient had no hair removal at all for certain operations or use of either depilatories or clippers rather than razors.

I can tell you, I don't have the data right in front of me, the national performance rates on the measure are extremely high. I am pretty sure Wanda has access to that data, and I know there has been some discussion about whether this is a, quote, "topped-out" measure or not.

CO-CHAIR TORCHIANA: Okay. Thank you, Dale.

It has been suggested to me that Laura Grosso from Yale University might be the developer for the other two CMS measures. Is she present?

(No response.)

All right, let's move on, then, to the next category, which would be AHRQ. Please cover all the potential measures on this list, if you can, whoever is here from AHRQ.

MR. BOTT: Yes. My name is John Bott. I work under contract onsite with AHRQ.

In one minute or less, yes, these

are a number of AHRQ quality indicators. They are derived from electronic administrative in-patient claims datasets.

And these are measures that are annually updated as far as coding, user input, refinements, to stay contemporary with the field of measurement. They are freely available to the public in terms of the software and the documentation, which feeds into users being able to provide continued input into the measure.

And real quickly, just to apologize for the late arrival of the citations to the measures, as we were in the process of updating the citations for all measures and we have since prioritized those in the NQF maintenance process. And we appreciate NQF's allowance for that late arrival of those updated cites.

And there are several others, when we talk about the RQIs, who are here in person as well as on the phone from the RQI team that

will probably be much more responsive to the more technical questions than myself.

CO-CHAIR TORCHIANA: Thank you.

I skipped Ingenix, the patients 18 years of age or older on a beta blocker. Is Laura Eaton on the phone from Ingenix?

MS. EATON: Yes, I am. Can you hear me?

CO-CHAIR TORCHIANA: We can hear you fine.

MS. EATON: Can you hear me?

CO-CHAIR TORCHIANA: We can hear you fine.

MS. EATON: Oh, okay. All right. Thank you.

Yes, I guess I will be short. I wanted to point out that we at Ingenix are aware of the STS registry measures that NQF currently endorses, and we believe that they are very good measures. Our hope is not to replace the STS measures, but to actually just be as a companion measure that would be added

value.

We feel that we created value in three important ways via the different data sources. We use claims data, and pharmacy claims data are highly good at capturing prescriptions filled. And we believe that our measure will be good at capturing patient beta blocker use.

We also feel that we look at a different aspect of care and measure a different part of the care process and communities. Since the STS measure uses registry data, it is supposed to be considered more provider-centric as when it is capturing more about physicians' prescribing behaviors. Our measure, it is a patient-centric which we are capturing information about a patient's prescription-filling behavior.

And, then, finally, our measure uses administrative claims data, as I mentioned previously. This type of data is routinely collected for reimbursement, and it

is not voluntarily reported, and, thus, could provide less-biased performance results.

Our measure also brings to the table health plans as a user. This administrative claims data is routinely collected by them, and it could benefit from using this measure.

I would also like to say that our measure does not require participating or using our proprietary database. Our proprietary database is used solely to test and create the measure. We put all of our NQF-endorsed measures out in the public domain for anyone to use and apply to their own databases.

At the last Subcommittee meeting, there was some question about the low CABG volume that we were picking up. And I just wanted to say that I went back and looked at the data.

There was a problem with some of the code, the programming code. It had to do

with we were trying to exclude CABG patients who were readmitted to an acute or non-acute care facility within seven days.

And it turns out that more than half of our qualified CABG members were being excluded from the analysis because of a bug in the code. So, it drastically brought up our numbers. Initially, we were reporting about 731 patients, and now that denominator is 2,303 patients. And so, we ended up with a new compliance rate of approximately 91.5 percent.

CO-CHAIR TORCHIANA: Thank you.

MS. EATON: I just wanted to make sure that I brought that to the Committee's attention because it was of great concern at the last meeting.

CO-CHAIR TORCHIANA: Thank you.

The next set of measures are from the ACS Quality Collaboration. We have two representatives here.

MS. SLOSBURG: Good morning.

I am Donna Slosburg. I am the Executive Director of the ASCQC, and this is Dr. David Shapiro with me.

I am going to take a little different track. I just wanted to make sure that he is aware about the ambulatory side.

The collaboration was formed in 2006 to develop standardized measures for the ASC specifically. It is a group of management companies, professional association, and accrediting bodies.

Something that is a little different is that the consensus was that our measures needed to be within the scope, influence, and control of the ASC. We didn't limit these to any particular patient population to allow all ASCs to participate.

I also wanted everyone to understand that we have tried to harmonize our measures. However, you need to take into account that ASCs are different than in-patient hospitals. The code set for billing,

hospitals use ICD-9; ASCs use CBTs. The claim format, hospitals use UBO-4s; ASCs use CMS-1500s. And our patient population is typically ASA 1 or 2. So, they are already, in our minds, risk-adjusted.

To date, just so everybody is aware, CMS has not implemented a quality reporting system for ASCs. So, all the work that we are doing is basically voluntary reporting. We also have public reporting on our website on our measures. And ASCs were not included in the EMR incentive program.

We are presenting three measures today for maintenance. Two are process measures, and they are in accordance with the surgical site prevention guidelines, IV antibiotic timing and appropriate surgical hair removal. I am sorry, the IV antibiotic is tomorrow. And, then, our third measure is hospital transfer admission. It is an outcomes measure.

We did harmonize the IV antibiotic

and the hair removal measure with the SCIP measures as best as we could. And we do have our results and our data from about 1300 ASCs on our website. It is out there for other ASCs to benchmark, and there's about 5200 ACSs in the country, but, again, these 1300 are doing this voluntarily.

David, do you have any comments?

MR. SHAPIRO: No. Thank you again for having us here to present these three measures for maintenance with you today.

I think Donna touched on the two probably most important aspects of why we would like you to consider our measures in possibly a slightly different light than you might others.

One is that they really are important to the ASC industry. This is a voluntary project that we have undertaken without to date any federal requirement for quality reporting, and we have done so with your help. And I want to, again, thank you

for endorsing the measures that you have in the past.

The measures really do warrant continued attention from us. Within our industry, they may not meet all the statistical criteria that we would like eventually to be able to apply to them by furthering our data-gathering techniques.

And the other thing is just a general comment on the ASC industry, which we are glad to talk more about in specific relation to our measures. But, in many instances, although we do comparable services, we are, for reasons including those that Donna Slosburg mentioned and others, we are comparable, but not identical to the HOPDs.

So, I think when we start talking about harmonization of our measures, we will be able to give you a little bit more background as to why we have been in some cases unable to harmonize to the extent that even we would like to harmonize our measures

with already-existing ones for other sites of service. So, I will wait until we discuss our measures, but, again, I want to thank you for your support in the past and hope that we continue to be able to work together on these measures and move forward.

Thanks.

CO-CHAIR TORCHIANA: Thank you.

The Society for Vascular Surgery? They will not be in until 1:30. Okay.

American College of Cardiology Foundation, followup assessment of stroke or death after carotid. They were going to be attending the meeting.

No one here from the ACC?

(No response.)

Okay. Children's Hospital of Pennsylvania, of Philadelphia, rather.

(No response.)

The same outcome.

I think that's it.

Any other developers that have

measures for consideration?

(No response.)

No. So, let's, then, go back to the agenda.

MS. GROSSO: Excuse me. This is Laura Grosso calling from Yale University.

They called for our measure a while ago. I am actually on the train. I am presenting on Measures 1550 and 1551. I am wondering, I was expecting to give the summary at 3:45. Would it be okay to give a brief summary when I am there at 3:45?

CO-CHAIR TORCHIANA: I think that would be fine.

MS. GROSSO: Okay. Great. Thank you.

CO-CHAIR TORCHIANA: Okay. So, let's now go back to the agenda, top of page 3, Measure 0127, preoperative beta blockade. We will start with the lead reviewer, Paula Graling.

MEMBER GRALING: Okay. I am

speaking for Work Group A, and we had just a few comments related to this measure.

It is a part of the STS CABG composite score. We felt that there was very strong evidence to support the measure and that we could see a clearly-demonstrated performance gap.

We did have some concerns related to the contraindications. And actually, STS provided us with an additional document that helps to update those concerns.

Feasibility was certainly still a question. We recognize that the cost of data extraction or a data extraction manager is present, as we have spoken about several of the other measures before. And we do recognize that there is some possible opportunity for harmonization as we consider all of the beta block measures that come before us.

And that's really it.

CO-CHAIR TORCHIANA: Any other

discussion on this measure?

(No response.)

Hearing none, if we could vote, first, on importance?

(Vote.)

Thank you.

I have to read it out. Twenty-one, yes; zero, no.

So, scientific acceptability. We now have a more complex vote here: 1 through 4, completely, partially, minimally, not at all.

(Vote.)

What is our numerator here? Is it 21?

If everyone could press again?

(Vote.)

Sixteen, completely; five, partially; no, minimally; no, not at all.

Usability, the same scale.

(Vote.)

We are still shy one vote.

Seventeen, completely; four, partially.

If I could ask the staff, who might be more knowledgeable in the technology here, if we just keep submitting votes until we hit 21, will that unload or overload or confuse the system, if everyone just keeps voting until we hit 21?

MS. WEBER: No, it should be fine.

CO-CHAIR TORCHIANA: Okay.

MS. WEBER: It only counts each vote once.

CO-CHAIR TORCHIANA: So, let's do that. Just keep hitting your votes until we hit 21. We might save 20 seconds a vote here.

So, feasibility?

(Vote.)

Seventeen, completely; four, partially.

Let's see if we can get it under 10 seconds this time and have a new record.

Does this measure meet all the NQF

criteria for endorsement? We are back to a 1, 2, 3 vote here.

(Vote.)

Wow! Okay. Are we done with that one?

Twenty-one, yes; zero, no.

Okay. So, we are done with that measure.

And, Paula, you have 0284?

MEMBER GRALING: Yes, that one, also.

Measure 0284 is surgery patients on beta blockers prior to admission who received a beta blocker during the perioperative period. This is a CMS SCIP measure, and you will recognize it. It is not just limited to CABG. It is, indeed, surgery patients.

Some of our original questions related to the timing of the perioperative period, and I think Dr. Bratzler discussed that in terms of looking at from day of

surgery to post-op day 2.

We also had a question about the denominator definition. You will notice in the definition that laparoscopy is identified in those lists, and we wondered why laparoscopes were just pulled out of that denominator. And perhaps he can address that for us.

Finally, there were some questions related to the collection methodology. It talked about getting the information from the medical record, and we wanted some clarity about paper versus electronic, or both.

And, then, finally, they did a nice job in terms of listing the disparities at the back of the application, but perhaps that they could address some of their disparity comments.

CO-CHAIR TORCHIANA: So, there was quite a bit of discussion about this measure on the call.

I will point out that this is only

for patients who were on beta blockers prior to admission. That confused me a little bit when I was thinking about it.

Is there any discussion that you would like to bring here?

MEMBER MORTON: Yes, I just wanted to chime in about the laparoscopic part of things. We have had this discussion previously about prophylaxis for VTE and everything else, but I think it just reflects kind of older ideas about laparoscopy, but, as we look now, one of the most common iterations is for bariatric surgery, where a lot of patients may be actually on beta blockade prior to surgery, hypertension occurring in about 70 percent of those patients.

So, I would highly suggest making sure that laparoscopy is not excluded from this, particularly vis-a-vis the bariatric surgery laparoscopic patients.

MS. JOHNSON: I believe Dr. Bratzler signed off. I can tell you -- this

is Wanda Johnson from OSMQ -- the laparoscope exclusion is being removed, starting with January 2012. And so, that exclusion will not be counted with this measure.

We do only collect chart-abstracted measures. Currently, we are in the process of updating specifications, so that they can be collected via EHRs.

MEMBER WILHOIT: One question that I had is the denominator is defined in terms of surgery patients on beta blocker therapy prior to arrival, but prior to arrival is not defined. Is that a year ago, a month ago, a day ago, a week ago? And I couldn't find a definition, and it seemed like it would make a big difference.

CO-CHAIR TORCHIANA: I think that was discussed on the call. Could one of the developers comment?

MS. JOHNSON: Sure. In the data element, beta blocker current medication, which is how the abstractors determine whether

it was considered a prior to arrival, it has to be listed as a home or a current medication. It must be documented that they were on it prior to arrival, that it wasn't started while they were in the hospital.

And if it is taken for a non-cardiac reason, then they would answer "no" to current medication. So, they wouldn't be in the denominator.

CO-CHAIR TORCHIANA: I guess I would throw in that the other discussion that occurred on the call was the vagueness of the criteria around what constituted a perioperative beta blocker, and I think that the vagueness in terms of what the dose, route of administration, et cetera, would mean.

I think the way that discussion was resolved was that greater specificity just resulted in greater complexity and measurement burden, and that just the threshold of was a beta blocker administered was the only way to do this simply.

MEMBER HALPERN: But she just said that it was for non-cardiac reasons. So, how is that determined?

MS. JOHNSON: It has to be documented that it was non-cardiac for them to answer no. That is just one of the outs that we allow them to have. Say they were taking it for migraines. It has to be documented that it was for a non-cardiac reason, taken at home, and, then, they could answer no to beta blocker current medication, and they wouldn't be in the measure denominator.

MEMBER HALPERN: So, what if they were just taking it for hypertension, because those patients often also have co-existing cardiac disease?

MS. JOHNSON: They would answer yes, and they would be considered for this measure.

MEMBER AFSAR-MANESH: Just to add on to what Dr. Torchiana was saying, one of the problems that I had with this measure was

that overall just having a beta blocker on in the peri-op period hasn't been shown to be beneficial if it is not titrated to a certain heart rate. And I understand that, for measurement purposes, that is going to be a lot more difficult to capture, but perhaps at a future date, as we move towards EHRs, the measure developers can think about how to include that. Because if we do want to improve the quality of care, just having metoprolol at a low dose on isn't really going to be beneficial if it is not titrated.

MS. JOHNSON: We will take the recommendations to the Technical Expert Panel. We do have an exclusion. As you know, they are bradycardiac, so the heart rate less than 50. Then, they would answer that there is a reason for not administrating a beta blocker and would get excluded.

But, yes, titration is something that we can address with the TEP.

CO-CHAIR TORCHIANA: Any other

discussion before we vote?

MEMBER WILHOIT: The one other thing would be a recommendation to -- some of the things that have just been discussed don't seem to be in the writeup in terms of excluding patients who are on a beta blocker for other than cardiac reasons, exclusion for bradycardia. Those are not listed as exclusions. It seems like those are pretty important and should be.

MS. JOHNSON: It is because they are listed in the data element, reasons for not administering. We will try to do a better job of getting those exclusions spelled out, instead of just listing the data element.

CO-CHAIR TORCHIANA: Okay. So, the first vote is importance to measure.

Remember, just keep pushing the button until it gets to 21.

(Vote.)

Did someone leave the room? I don't think so.

Twenty-one in favor.

The second vote is scientific acceptability. Here we are voting on the 1-to-4 scale.

(Vote.)

Ten, completely; ten, partially; one, minimally.

Usability, on the 1-to-4 scale.

(Vote.)

Twelve, completely; nine, partially.

Feasibility, on the 1-to-4 scale.

(Vote.)

There is definitely one in here that is not working.

Twelve, completely; nine, partially.

Does the measure meet all the NQF criteria for endorsement, yes, no, or abstain?

(Vote.)

Nineteen, yes; two, no.

Okay. The next one, beta block at

discharge. Dr. Stafford?

MEMBER STAFFORD: Yes, that's mine.

This is beta blockade at discharge and isolated CABG, just so everybody understands that that's what the measure was about.

In terms of importance to measure, everybody in the group really felt that it was important to measure, that there is clearly a performance gap. But even though the mean was high, the median showed that a fair number were off the scale and still needed to be measured. So, we all felt it was important to measure.

In the discussion about scientific acceptability, there were some questions about patients with contraindications who were removed from both the numerator and denominator, and, also, the time windows. And those were pretty much cleared up in the phone call with the developer. I don't think there

were any big issues that I can recall.

Everybody felt that for the most part usability was not much of an issue.

And feasibility, there were questions about cost for data abstraction that were answered.

Lots of questions, it got around, again, to gaming the system and whatnot, but pretty much these are things that are going to occur with a lot of the measures that we have. Everybody felt that this was at least a measure that the outcome measure was relatively related to the process measure. And so, from that standpoint, we felt it was a good measure to have.

CO-CHAIR TORCHIANA: Comments or further discussion?

MS. MURPHY: There was a question in the Work Group discussion that the group I thought felt Dr. Shahian could answer about why no risk adjustment.

MR. SHAHIAN: Well, I think for

process measures, typically, exclusions and eligibility requirements generally take the place of risk adjustment. I think it would be difficult to construct a risk-adjustment model for this, although some people have advocated using risk adjustment for process measures, but I think, in general, eligibility and exclusion criteria have been used instead.

CO-CHAIR TORCHIANA: Okay. Let's vote on this.

Importance to measure, yes, no, or abstain.

(Vote.)

Twenty-one in favor.

Scientific acceptability, on the 1-to-4 scale?

(Vote.)

Eighteen, completely; three, partially.

If I could ask, again, a technical question, do you have hit Send every time you hit the number if you are voting repeatedly?

MS. WEBER: You have to hit Send.

CO-CHAIR TORCHIANA: Okay. So, it takes two fingers to do this to keep it going. Hit 1, 2, 3, or 4 and Send until you see 21 up there, and maybe we will get under 10 seconds consistently then.

Usability, a scale of 1 to 4.

(Vote.)

Seventeen, completely; four, partially.

Feasibility, scale of 1 to 4.

(Vote.)

Eighteen, completely; three, partially.

Does the measure meet the NQF criteria for endorsement, yes, no, or abstain?

(Vote.)

Okay. Twenty-one, yes.

Where is the antenna actually located that is receiving these? It is on your computer? Okay. We can't ask to put you in the middle of the room. I think that would

be unfair.

(Laughter.)

Okay. The next measure is 1480, the Ingenix measure that we heard about, patients 18 years of older.

Dr. Stafford again.

MEMBER STAFFORD: Yes, this measure led to a lot more discussion. Again, I think everybody said, well, isn't this really related to the previous measure, and should this somehow be harmonized, was one of the questions that did come up for the group.

Everybody felt, again, that it was important to measure beta blocker use in this patient population. So, from an importance standpoint, everybody felt that it was reasonable.

As to the scientific acceptability, there were lots of questions. Some of those were addressed today about the population, and particularly, the numbers of CABGs that were actually included in the

group. Clearly, that was something that we picked up on, and they found that, actually, there was a bug in their computer system.

It still isn't clear to me, at least hearing what we got from the developer today, whether those numbers are still high enough for us, at least for me thinking about it.

There were, similarly, other discussions about this being a proprietary database and, therefore, it only applies to people that actually have insurance benefits. And so, there may be a large number of patients in the population, particularly in the Medicare and Medicaid population, where this may not apply. And so, it may be not be generalizable to the large population that we would like to see the measure looked at.

The flip side of that discussion was that, for health plans, it would be usable because it was another way for them to get that information.

Those were really the two biggest things about the database.

As to the feasibility, it varied from minimally to completely. And again, it depended on what patient population people wanted to look at.

CO-CHAIR TORCHIANA: Other comments?

MEMBER CIMA: How is it going to be attributed?

MEMBER STAFFORD: What was the question?

MEMBER CIMA: Well, whose scorecard is this going to go on, the hospital's, the physician's? I mean the cardiac surgeon can give a prescription for a beta blocker. The hospital can give a prescription for a beta blocker. The health system can give a prescription for a beta blocker. But if the patient doesn't fill the prescription for the beta blocker, where is this data, who is going to make the

performance measure? Who is going to improve it? Who is the audience to change to get quality?

MEMBER STAFFORD: Exactly, and there was a lot of discussion about that, whether this would be used for internal benchmarking in health plans versus external benchmarking; how would people use that?

And it was interesting because, originally, we all thought that it meant people would actually have filled the prescription. But, after further discussion, it was actually just that there was a prescription written. They didn't actually have to have gotten it filled. So, it doesn't even really get to what we would all like to measure: did they actually get the drug and did they actually take it?

MEMBER CIMA: So, in that case, isn't it the same as prescription at discharge?

MEMBER STAFFORD: Yes, it is just

they are looking at a different population.

MS. JOHNSON: This is the developer. Could I just make a comment?

Yes, when we spoke last time, we did talk about codes that for prescriptions written. However, there is no problem with us taking out those codes and have it only strictly represent pharmacy claims that have been filled.

We just initially put them in there because we were trying to make it as consistent with STS as possible.

CO-CHAIR TORCHIANA: I think what I would say, the other issue that came up on the call in response to this line of question was the fact that this comes off a different dataset and publicly-accessible dataset was seen as one of the selling points and for complementary values of this measure above and beyond the STS measure.

So, the way that I sort of walked away in thinking about this was that it

represented a different look from a different perspective in a different population, that it potentially had the value that it could go to claims or prescriptions that were actually filled, but that it had the disadvantage that it was a very narrow subset of the total CABG population, which is more than half Medicare, and that it at least originally was probably not terribly accurate in identifying those patients. It sounds like that accuracy has been tightened up, although I agree with Dr. Stafford; I am not exactly sure that we know to what degree it has been tightened up.

MEMBER HALPERN: I would also say, if you just look at prescriptions filled, since a lot of these patients come in on beta blockers, they may go back to taking their own beta blockers and not fill a prescription afterwards. So, that doesn't mean that they are not taking it.

MEMBER WILHOIT: Have we seen the revised data? I think seeing the numbers is

important to understanding what this measure is getting at.

MS. JOHNSON: You are referring to the new denominator that I talked about earlier?

MEMBER WILHOIT: Yes. Well, the new denominator and, then, the rate.

MS. JOHNSON: Yes.

MEMBER WILHOIT: All of the results that we had seen originally apparently aren't valid. So, what does the data show?

MS. JOHNSON: Okay. No, I haven't submitted that. I definitely can.

What we did find was that there was about 2300 patients in the denominator and 2100 patients in the numerator, giving us a compliance rate of 91.5 percent.

MEMBER HALPERN: I also have a question that I didn't see specifically. It is, how are you defining that a prescription was given? Is it just that it is on the going-out medication reconciliation, or how is

it being defined?

MS. JOHNSON: It was getting defined through CPT codes, beta blocker therapy prescribed, which I believe is where there is success. That is what it was going out after. But that is the code that I suggested that we can remove from it and just have it strictly be pharmacy claims data, outpatient pharmacy claims data.

MEMBER CIMA: So, if it is based on outpatient pharmacy claims data, then that means that they had to fill it?

MS. JOHNSON: Right.

MEMBER CIMA: Okay.

MEMBER HALPERN: So, I go back to my original point. It is that patients may already be on beta blockers and, then, they will not be filling a new prescription because they already have medications, certainly not within seven days. They may fill it later on, but, then, you are going to ding people for things that may not be an issue.

MS. JOHNSON: Yes, you are right, though our compliance rate went up. I guess the bottom line is that you are right; that is something to consider.

DR. BURSTIN: Just one clarification. Actually, in the measure submission form -- this is Helen Burstin -- it specifically says that you have to fulfill at least one of the three criteria, one of which is you filled the prescription. So, it is prescription filled --

MS. JOHNSON: Right.

DR. BURSTIN: -- within seven days after hospitalization.

Charge two, that you had a claim for that, and a beta blocker prescribed in the 35 days prior. So, that might answer a part of your concerns.

And, three, that the patient had a claim with a procedure code of beta blocker at discharge. So, that is the only one she is mentioning, but, in fact, there are two others

that she is not talking about that I think address what you are saying.

MEMBER WILHOIT: It sounds like a number of things have changed from what -- I think we were just told that the HCPCS codes were no longer, but at this point I am confused about what is in the measure, what has changed, what the results are. I am not quite sure what we have got.

MEMBER STAFFORD: I would have to agree with that. And if I were to vote right now on this measure as it stands without all of the other qualifications that we have heard, I would say it is not important to measure.

So, yes, I would request that we either table this and come back and rediscuss it or we vote on the measure as it stands.

MS. MURPHY: So, if you are looking at tabling it for discussion at a later time, what you need to do is say, in addition to what you have just said in terms

of information that you need, in order to consider, to properly consider, the measure, and so, you could table it to do that.

There were a couple of other things that during the Work Group were discussed that I don't think you have mentioned. One was some question about the adequacy of the validity testing and whether or not you were comfortable with that at this point.

And another one I think with respect to the plan for public reporting, there was a question about that in terms of the extent to which public reporting would occur.

So, I just saying that, so if there are conditions that you want to apply for later consideration, it is probably time to articulate all of them.

MEMBER WILHOIT: I think I am the one that raised the issues about validity testing. And for those of you that didn't go

through all the fine print, the validity testing related to public reporting consisted of surveying two customers. Well, customers aren't the public, and two isn't a very big number. So, I certainly was not comfortable that that criterion was met.

MEMBER STAFFORD: Dr. Dutton commented on the validity issues as well.

MEMBER CIMA: If we table it, we are somehow saying that it seems to be a useful measure and we want to bring it back to make it robust and better. My fundamental thought is I don't think this is a useful measure. So, why don't we just put it out of its misery?

(Laughter.)

MEMBER FINDLAY: For clarification, this is not a publicly-reported measure. Or is it? Yes, that's what I thought. I was just confused by that discussion there.

CO-CHAIR TORCHIANA: So, let's

have a little further discussion on the table versus vote it up or down.

I guess the vote up or down would be just to say we will keep it alive for further consideration versus we don't think is under any circumstance going to lead a valuable new measure. That vote, obviously, isn't up on the slides, but we could do that perhaps by a show of hands, if that is an acceptable process step to take.

MEMBER COLLINS: And from what I remember, this is very similar to the lipid-lowering agent proposal at the last meeting by the same company, which I think we did, quote, "kind of put out of its misery", if I am right. We could clarify that, though.

MS. MURPHY: So, the critical question is, from what you have heard and from the additional information that you have gotten in terms of changes and updates, would it be useful for you to see that additional and/or corrected information prior to going to

the vote about the importance of the measure? Or do you feel confident that you understand what that is and comfortable with proceeding to a vote without that?

And I would say that you would really only be able to go beyond importance if you got the additional and corrected information. Otherwise, you would not be voting on current specifications.

MEMBER ROGERS: I think this issue highlights a problem we have with some other measures that we will look at, which really has to do with followup and where patients go after they leave the hospital environment where the records are easily available.

And although this may not be a perfect measure, it does help us address that period of time after the acute event has occurred, which is a challenge on all long-term followups.

I mean Ingenix is a large database. It is not comprehensive. It

doesn't have everybody, but it is a very good start to answer questions that I think are important.

And in the context of what they do, it probably works reasonably well. So, I just think we have to keep that in mind because long-term followup is a challenge that we are going to face on all of our measures. It is one thing to do something in a hospital; it is quite another to know what happens to those patients a week, a month, or six months later. And this is a step in that direction.

So, I think that we shouldn't dismiss it out of hand. It may not be perfect, but it is I think valuable.

MEMBER SAIGAL: While I would agree with you on some of those points, one of the sidebar discussions we had earlier out in the hallway was I think understanding the impact of what putting the NQF stamp on these measures has long-term for people.

So, by putting this stamp on a

measure, that gives it some street cred elsewhere with groups like CMS or JCAHO to pick up that measure and, then, require that people use it.

So, I think you have to be careful, and we really need to think about what we are doing when we put an endorsement on a measure. NQF endorses the measures. They don't monitor them. In terms of implementation, they don't force implementation. But the fact of the matter is that is the effect that the NQF endorsement has.

I think we need to be really careful about endorsing measures that really don't broadly cover a large population. I understand your point about the long-term followup. I am not sure this is the best way to do it.

MEMBER ROGERS: It may not be, but it also cuts both ways. And this relates to a discussion that I had prompted at the last

meeting, having NQF approve measures which actually, if not toothless, have less value in the context of our ability to get more information.

So, we have just approved, for instance, mortality measures in cardiac surgery, when, in fact, that may not be a particularly useful measure and there is more information available which would be valuable.

So, NQF has just put its or will put its stamp on something for the next three years which will be, I think, woefully behind the times. So, it cuts both ways.

MEMBER WILHOIT: Coming from the perspective of a health plan, having administrative measures, being able to measure the care provided for the health plan members by health plan providers is really useful.

What makes me so uncomfortable about this measure is there are 15 million members in the database. I believe that Ingenix said that there were 2200 members

identified with bypass surgery. Now this measure is limited, I believe, to members of the pharmacy benefit. I am going to make a rough guess that 5 million of the people have a pharmacy benefit.

So, that is around 400 people per million, which, compared to our data -- and I polled our data to see -- is probably about, for our commercial population under age 65, which matches up with this measure, looks to me like maybe they are finding half the people that I would expect to have had a bypass. So, again, the numbers make a difference.

Also, if you are finding 500 patients per million, what is the use of the measure? Are we looking at the health plan level? At the health plan level, 500 members is okay if I have a million members.

If I am breaking that down by hospital or by doctor, you end up with numbers that are so small that you somehow add credibility to I am reporting on your 13

members who were in my health plan and had a pharmacy benefit, and I am reporting for another doctor on his 23 members. And that isn't useful.

That is the thing that bothers me, is that if you break it down by provider, the numbers won't be big enough to be useful.

CO-CHAIR TORCHIANA: So, I am going to take the prerogative of calling the question here, since we are over time. So, I think we can address this, actually, with the vote on importance. So, if we vote importance up, then we will proceed to the other criteria and ask that the measure be optimized. If we vote importance down, I think the topic is closed for today at least.

So, let's do that vote. It is a yes or a no.

(Vote.)

So, I think that is a fairly definitive no on that question. So, we will wrap up the voting on that item.

Now, before we break for lunch, we have a required public comment. Should we hold that for later, Melinda, or should we do that prior to lunch?

MS. MURPHY: We should do it now.

CO-CHAIR TORCHIANA: Okay. NQF member/public comment, those on the phone or in the room?

(No response.)

Hearing none, we will go to lunch. Lunch break is one-half an hour. Lunch is outside.

(Whereupon, the foregoing matter went off the record at 12:41 p.m. and resumed at 1:11 p.m.)

A-F-T-E-R-N-O-O-N S-E-S-S-I-O-N

1:11 p.m.

MEMBER DUTTON: All right. Pancreatic resection mortality and volume. The group generally thought these measures were important, usable, feasible, and met scientific applicability.

The particular issues identified were, first, discussion about whether there should be two different measures or one harmonized measure or one measure that combined both. I think we have had some of that discussion about esophageal resection already, and I think probably the same would apply.

It is two measures, but a companion or paired, so they would always be reported together, probably makes the most sense. It is both simple and gets both pieces of data on the table.

There were a couple of questions about harmonizing just these two together.

They actually measure slightly different things, in that, the mortality measure, is specifically for cancer surgery; whereas, the other measure is not. It is for any pancreatic resection.

The volume measure, looking at the ICD-9 codes, it is only complete pancreatic resection. So, we sort of wondered what happened on the partial resections and partial operations. One of them listed a lot more codes than the other. That should probably all be harmonized.

There was a problem. I identified that the mortality measure, actually, both measures exclude transfers. So, if you transfer the patient to another acute care hospital, they disappear from both the numerator and the denominator, which, if you are a small volume pancreatic hospital and you are getting bad results, if you can ship the patient downtown, this seems like a great racket. And I know that happens. So, that

would seem a bad exclusion there.

And, then, finally, there is one section with the volume one where it was actually, it said esophageal; it is just a typo or a missed cut and paste from the other measure that they need to clean up.

I think that was everything that we identified as the important issues.

CO-CHAIR TORCHIANA: Could we have a comment from AHRQ on why transfers are excluded?

MR. ROMANO: Yes. This is Patrick Romano. I am happy to join you. I am a general internist supporting the AHRQ quality indicators program from UC Davis.

So, just to address a couple of those questions, transfers out are excluded because, of course, the outcome of those patients is unknown if you only have data from your own center or if you are using a dataset, as most users do, that is not linked across sites.

So, ideally, of course, we would like to have a 30-day measure, and some payers would be able to specify a 30-day measure. But the AHRQ measures are designed for applicability to a broader range of datasets which do not generally permit that kind of linkage across hospitals.

We have looked at this empirically in the California dataset that does allow linkage. And you might be surprised to find that, in fact the post-transfer mortality is quite low because of those are sort of back-transfers to a community hospital setting or a rural hospital after the patient has had the aggressive procedure at a teaching center. And, then, they go back to a site that is closer to their home for some further recovery. So, in fact, what we see empirically is that post-transfer mortality is quite low in the linked dataset in California, anyway.

The other issue as far as the

difference in the denominators, this is an issue that we have discussed in the first round of endorsement. It has also been discussed with Leapfrog.

The rationale, basically, is that, of course, the mortality is much higher for the cancer surgery, but the benign surgery, experience with benign surgery appears to translate into better outcomes for both patients with benign disease as well as patients with malignant disease.

So, including benign disease in the volume measure kind of gives providers the benefit of the doubt for doing procedures on patients who have benign disease. So, that is the rationale.

But it is subject to change. It has just evolved through discussions with expert panels previously.

CO-CHAIR TORCHIANA: Any further discussion?

MEMBER STAFFORD: Yes. Hi. I am

Renae Stafford, and I am actually the one brought up why the benign patients were excluded. Can you give me a little more background on that? I still don't quite understand the rationale.

MR. ROMANO: Well, these indicators are based on an empirical literature that is overwhelmingly based on cancer surgery. So, there have been a couple of large meta-analyses, mostly based on observational studies, and they were limited to pancreatic cancer surgery.

There is, of course, a difference in mortality between patients with pancreatic cancer surgery and patients with surgery for benign disease. That difference in mortality could theoretically be adjusted for through risk adjustment. So, that is an option. It is just that we kind of tailored the original definition to what other investigators have reported on in the literature.

MEMBER HALPERN: So, should it be

titled, instead, pancreatic resection for cancer? Because it seems to me, actually, it would be more important to know what the mortality is for benign disease since the patients wouldn't otherwise necessarily die from their disease process.

MR. ROMANO: It is a valid point. At this point, you know, the indicator specification is what it is, but we can certainly perform some additional analyses to explore what the impact would be of including or stratifying separately for benign disease.

MEMBER DUTTON: My recommendation there would be just to count it, you know, all pancreatic disease in both the numerator and the denominator, both of these measures. It would keep them harmonized. It is a low-enough-volume condition that I think trying to subset it further isn't really helping you.

MEMBER HALPERN: I actually was more implying the measure would be more for benign disease, a separate measure. Because

I think measuring mortality for benign disease, as I said, is almost more important because you wouldn't otherwise die of your disease process, and you would otherwise die of pancreatic cancer, if you don't have your resection.

MEMBER DILLON: Right. I agree. I think you have to split these two out.

MEMBER CIMA: Other than sort of emergency sort of pancreatic debreedment, what other benign conditions lead people to have major pancreatic surgery? I mean other than IPMNs, but that is now almost becoming classified as a malignancy.

So, the question is -- I think the point is that it's such a low-frequency thing. If anything, I would say exclusion should be a diagnosis of pancreatitis. I mean those patients have a 40 percent mortality or 30 percent mortality as is.

But, you know, what benign condition do we operate on in such high

volumes of the pancreas?

MEMBER HALPERN: The volume is a paired measure, is that right? Yes, so they are always reported together.

MS. MURPHY: That is the way they were endorsed, to be reported as paired measures.

MEMBER DILLON: I am just looking through the exclusion criteria, again, just to make sure that, you know, in cases of splenectomies and the tail of the pancreas gets removed, I am just worried about the possibility of coding variability here that could cause a problem with the data.

A resection for a renal, you know, a major surgical oncology procedure in which you do a partial pancreatectomy but it is for a primary renal or splenic tumor.

So, they would have to be able to exclude those as well.

CO-CHAIR TORCHIANA: Could I ask Melinda for some guidance? We have some

suggestions of some fairly-significant modifications in the way that this measure is constituted. Do we just proceed and vote on it as is, or how would you suggest we go ahead?

MS. MURPHY: In much the same way you have done, which would be, if there are recommendations, suggestions, conditions that you would apply to the measure, that those be stated and that we get the response from the developer, and, then, vote the measure based on that.

CO-CHAIR TORCHIANA: Well, I would suggest that we have heard enough concerns around the table over inclusions, exclusions, how to address this, that we are not prepared to vote.

So, should we put off the vote and lay out what those concerns are? I think pancreatitis is clearly one. Resection for pancreatitis probably shouldn't count as a pancreatic resection for this purpose.

And, then, the question as to whether or not we should include and lump benign and cancer or whether we should split and/or exclude benign from one or both of the measures.

Any further comments on those?

MEMBER DILLON: Let me just sort of muddy the waters just a little bit, though. And I will aim this at sort of Vivienne and Robert and the other surgeons in the group here.

In terms of the impact of disease, which is what you are concerned about, within a 30-day window, whether we are dealing with pancreatic CA in nodes or a small cyst or an IPMN in the tail of pancreas probably doesn't matter, which, as I talk about it and think about it, brings us back to a more unified approach, you know, as this proposes.

Because I like this measure, but --

MEMBER HALPERN: But to include

all pancreatectomies?

MEMBER DILLON: Yes.

CO-CHAIR TORCHIANA: So, that would include all pancreatectomies for malignant or benign disease for both mortality and volume with the sole exclusion of pancreatitis?

MEMBER DILLON: Yes, I think so.

MEMBER WILHOIT: One of the things that would be helpful there would be to look at the numbers. In one of the measures, in 0366, at the beginning under disparities, it gives some volumes for a sample. But in 0365, it doesn't give the volume. It just gives the rates.

So, the numbers are probably available to look at. It would be interesting to see how the volumes compare for benign and malignant disease and, also, how the mortality rates compare.

MEMBER ROGERS: I think it makes sense to kind of keep them together for any

reason that these are relatively rare operations. There's not a lot of them going on. I think there were maybe 3,000 Whipples done last year. So, there's not a lot of them. So, that is one argument to kind of keep them together.

The other thing is it is mainly the organ that helps determine a lot of these complications. So, regardless if there is cancer or if there is any other reason for the operation, it is the fact that the pancreas is so hard to work with that it can lead to problems.

MEMBER SIPERSTEIN: We actually had a very similar discussion about the esophageal measure last time, where the mortality figure dealt with cancer cases, and the volume credited the hospital with a few extra percent for overall esophageal volume, but I don't think it really is changing the way the wind is blowing on this because the number of pancreas procedures for benign

disease is relatively small. If a hospital, say, upgrades their numbers by 10 or 20 percent, by including all pancreatic procedures, it is not going to really change the impact of the measure one way or the other.

So, it may be a lot of discussion about something that really isn't affecting the bottom line on this.

MR. ROMANO: Yes. This is Patrick Romano again.

I just wanted to emphasize that these measures were created through a process of reviewing the published evidence. So, again, the published evidence is limited to, largely limited to resection for pancreatic cancer because that is the overwhelming majority of the pancreatic resections.

So, I appreciate the point about benign disease, but it is kind of going beyond the evidence on which the indicator was based. So, that is why we have an Expert Panel, we

have an AHRQ Expert Panel process that led to this specification. The decision was made to give the hospitals credit, essentially, for the additional volume, that they might get 10 percent, 20 percent additional volume with benign cases. But the core of the concept is based on the literature on pancreatic surgery and mortality.

With respect to exclusions, that is easier to address. I think we could easily accept exclusions that this group wants to propose as just part of the next revision process.

CO-CHAIR TORCHIANA: And so, I would say we table the vote for the time being, that we ask for data on these topics that you just explained, and to contemplate an exclusion for pancreatitis, and that on further review a decision ought to be made whether to try to harmonize the volume and resection mortality all into the same boat.

So, are we okay to move on to the

next measure?

MS. MURPHY: And we would need to bring that back by the time of the call that we will have in follow up to this meeting.

CO-CHAIR TORCHIANA: Okay. All right. So, it is now 1:27. I don't think we will get through perforated appendix in three minutes.

Do we have our vascular surgery developers on the phone?

MEMBER SEARS: Nick Sears is here.

CO-CHAIR TORCHIANA: I'm sorry, could you say that again? I wasn't able to understand.

MEMBER SEARS: Sure. No problem. This is Nick Sears.

CO-CHAIR TORCHIANA: Great. Very well.

So, there are a series of vascular surgery measures before us for consideration. Could you give us a brief summary, as the developer?

MEMBER SEARS: I wasn't the developer.

CO-CHAIR TORCHIANA: Oh, I'm sorry.

MEMBER SEARS: I was the reviewer.

CO-CHAIR TORCHIANA: I thought you were the SVS person. I apologize.

We need to wait for the Society of Vascular Surgery person to come on the phone.

So, is there any way that we can be notified of that, other than just asking?

MS. MURPHY: Yes. If you want to go ahead to the next measure, and, then, we can check --

CO-CHAIR TORCHIANA: At the end of that?

MS. MURPHY: -- when that is completed.

CO-CHAIR TORCHIANA: Okay. All right. Very good.

So, we are at perforated appendix.

Dr. Morton?

MEMBER MORTON: Yes. This is another AHRQ measure. It is perforated appendix admission rate.

This has been something that has been measured quite frequently in surgery from time immemorial to know if it there is a quality problem.

The particular measure is important for a few reasons. It has got scientific validity. It has been very easy to implement. Unlike a lot of the administrative measures that are out there, this has actually had a referendum clinically where it went into patients' charts to confirm the actual diagnoses. Dr. Romano was involved with all of that. So, from that end, it is pretty useful.

In terms of meeting some of the measures, it is useful in looking at different things such as clinical management, waiting until an appendix perforates can have some untoward consequences.

There is also a big performance gap. I was surprised to see it is on the order of about 20 to 30 percent of appies are perforated. So, there is still a lot of potential opportunity there.

A lot of data in the literature around the topic, particularly when it comes to access to care. It seems like there is a higher rate based on your socioeconomic status, based on your rural status. So, there is some evidence for disparities that haven't been addressed.

It has some pretty important consequences when there is perforation in terms of longer length of stay and other resource utilization.

As I mentioned before, there is quite a bit of, quite a few perforated appendectomies that are out there.

So, those were the main points. It seems to be a feasible and important measure with scientific validity.

CO-CHAIR TORCHIANA: Any other comments?

(No response.)

Okay. Oh, I'm sorry.

MR. ROMANO: I just wanted to clarify for the group that this is a population-based measure. So, it is not a measure that is intended for attribution to individual surgeons or hospitals.

CO-CHAIR TORCHIANA: Right. We discussed on a call it is reported by region.

MR. ROMANO: Right.

CO-CHAIR TORCHIANA: Okay. So, let's go to the vote.

Importance to measure, yes or no.

(Vote.)

Nineteen, yes; two, no.

MR. KRESOWIK: Hello. I don't know if you can hear me or not. This is Tim Kresowik from the Society for Vascular Surgery. I was on before, but, apparently, somehow my line was muted, so I had to go

through the operator to get in. So, I am here; I was here, and still am.

CO-CHAIR TORCHIANA: Okay. We will be with you in about 30 seconds. We have gotten really quick at voting these in.

(Laughter.)

So, we have just three more votes. It should take us about 30 seconds.

Scientific acceptability, on a scale of 1 to 4.

(Vote.)

Keep pushing those buttons.

Sixteen, completely; five, partially.

Usability.

(Vote.)

Eighteen, completely; two, partially; one, not at all.

Feasibility.

(Vote.)

Eighteen, completely; three, partially.

Finally, yes, no, or abstain, does the measure meet all the NQF criteria for endorsement?

(Vote.)

Twenty, yes; one, no.

Thank you.

So, if we could hear, then, on the SVS measures, please, from the developer?

MR. KRESOWIK: Thank you.

Again, this is Tim Kresowik. I am a vascular surgeon at the University of Iowa, representing the Society for Vascular Surgery.

I have been involved in measure development going back a long ways, back to the original Cooperative Cardiovascular Project, and have spent a long time acting as a consultant, a facilitator for all kinds of measure development.

So, I would, if I could, try to introduce a group of measures, just sort of as an overview, because I think these are kind of a unique approach. So, I would like, if I

could, just to start out and talk about Measures 1523, 1534, 1540, and 1543.

CO-CHAIR TORCHIANA: That is exactly how we would like to do it, too, if you could cover them all together.

MR. KRESOWIK: Okay. Perfect.

So that, the basic concept here is to try to get away from some of the problems and perverse incentives that are associated with current approaches to measurement of, if you will, surgical mortality or other complications.

What we are trying to do with this group of measures is to try to focus on a procedure and an outcome that requires minimal risk adjustment. And in this cases, what we are talking about, I will start first with probably the carotid procedures would be easier to introduce.

But the basic principle here is that the majority of carotid revascularization procedures performed in this country are done

on asymptomatic individuals. And in that setting, the risk versus benefit is quite narrow. So, it is essential for those groups of patients that the procedures are performed with an extremely-low mortality/morbidity. And an important component of that is the selection of patients who are at low risk.

And that becomes part of the surgical decisionmaking and, therefore, requires no further risk adjustment. In fact, you would not want to, if you will, give credit for someone performing interventions on someone who is at high risk because of comorbidities or people who have shortened life expectancy because of their comorbidities. Because in order to achieve benefit in that patient population, the asymptomatic population, the risk needs to be quite low.

And it is always hard; I have participated in many of these on a phone versus an in-person meeting. So, I am just

going to stop there for a second and just see if everybody is still with me or if there are any questions about that overall principle.

CO-CHAIR TORCHIANA: Any questions?

(No response.)

That sounds like a very succinct approach.

MEMBER SEARS: I have one question. When you say asymptomatic, so would you qualify somebody who has a transient ischemic attack but, otherwise, then resolves and doesn't have another one, is that an asymptomatic patient or is that a symptomatic patient at that point?

MR. KRESOWIK: A very important question. And the issue in this case is we actually talk about three levels of symptomatic status. And again, I don't know the composition of the group in the room. So, I am going to try to keep it at a fairly, even, if you will, a layperson level.

But the issue as far as the data, the randomized trials that have been done support a strong benefit for patients who have hemispheric or clear strokes or transient ischemic attacks within a relatively short period of time. The studies, let's say, for the major trials are approximately 120 days or four months.

So, for the purpose of the way we have defined this measure, those measures would be classified as symptomatic. So, these are patients who have clear hemispheric that is on the same side of the carotid that is being operated on within four months of the procedure.

We would have a second group which we would call other symptomatic, which would include the patients that would have had an event sometime in the remote past, okay, that is beyond the 120-day window, and would also cover patients who have neurologic events in the vertebrobasilar circulation or perhaps

have had events in the contralateral hemisphere to the carotid being operated on.

And so, those two groups would be separately classified. But this measure is focusing on everybody else, which is the patients who are truly asymptomatic in the sense that they have never had a stroke or transient ischemic attack in any distribution at any time. So, these are purely asymptomatic.

And again, from good data in the Medicare population, that is still approximately 60 to 70 percent of the patients who are receiving at least for carotid endarterectomy in this country. So, it is the vast majority, but these are patients who are completely asymptomatic.

Does that answer the question?

MEMBER SEARS: Yes, it does. This is Nick Sears. I am a retired vascular surgeon.

So, it is just logic there, yes.

MR. KRESOWIK: Okay. So, I note there are some other measures. I will just talk about I have been one who has strongly advocated against volume-based measures because they, in fact, encourage overuse and potentially, as I said, create a perverse incentive for people to operate on more and more patients for whom the benefit is small.

So, I think the other advantage of this approach is it does not have the perverse incentive of encouraging overuse, and, hopefully, in that way, also, will lead to cost reduction, if you will, in the country.

So, the advantage of this approach, as I said, is the simplicity of, once you define this population as the asymptomatic, you need to do no further risk adjustment. And, then, you can focus just on the event, which in this case would be stroke or mortality.

I can stop there, and, then, I can go on to the aneurysm measures, but --

CO-CHAIR TORCHIANA: No, please go on.

MR. KRESOWIK: Okay. So, do you want me to talk about the aneurysm measures then?

CO-CHAIR TORCHIANA: Yes.

MR. KRESOWIK: So, the aneurysm measures are designed to be the same basic principle. Again, the data that would support aneurysm repair in the case of non-ruptured aneurysms, this is purely a preventive operation. So, the risk versus the benefit has to be balanced on an individual patient level.

We have pretty good data about the likelihood-of-rupture risk based on size. And there is a slight size difference between men and women, but the basic principle here is, if you focus on the non-ruptured aortic aneurysms that are, if you will, of small to what we would call moderate size, you don't need to do further risk adjustment, just as I discussed

with the carotids.

The basic principle here is that that should be part of the decisionmaking. And you don't, again, want to give credit to someone who is taking a very sick individual or someone with a shortened life expectancy and fixing a small or moderate-size aortic aneurysm. The likelihood of harm versus benefit in that case would weigh against the procedure.

So, it is exactly the same as what we were talking about before. It minimizes the need for any further risk adjustment as it is part of the surgical decisionmaking.

So, I will just stop there. That is really the principle before these four measures. It is just separating into separate categories the endovascular approach to aortic aneurysms versus the open approach, the two separate measures focusing on those populations, and the same for carotid endarterectomy and carotid stenting.

CO-CHAIR TORCHIANA: Do you want to say anything about the surveillance measure?

MR. KRESOWIK: Sure. The surveillance measure is, again, based on the principle that, if the patients are undergoing endovascular aortic aneurysm repair, they do require some post-procedure surveillance to make sure that there are no potential complications that could increase the risk of future rupture.

There is, you know, I think a fair amount of controversy as to exactly what the time interval needs to be, exactly what the imaging modalities need to be. This measure is designed to be, it is sort of, I guess, a very minimalistic approach of saying at least one surveillance study between 3 and 15 months, which is really sort of the floor, if you will, for what might be ideal and does not necessarily specify the type of imaging study, just that that imaging study needs to be one

for which you can comment on size and any evidence of endo leak.

So, again, I don't think anyone would argue that this would be the absolute minimum that should be done for these patients.

CO-CHAIR TORCHIANA: Okay. Thank you.

MEMBER MORTON: I had a question, if I could? The question is, given that the endo leak rate is so high -- it has been estimated 30 to 60 percent in some studies -- should you be more prescriptive about the frequency of imaging as well as the type of imaging, since this appears to be a very common occurrence?

CO-CHAIR TORCHIANA: Why don't we get to that when we discuss?

Tim, are you going to be able to stay on the call or are you --

MR. KRESOWIK: No, I certainly can. I am available.

CO-CHAIR TORCHIANA: Okay. So, why don't we get to that when we get to specifically discussing the measures?

So, thank you very much for that introduction. We are now going to go back to our agenda and follow along in the order that we were set to go.

So, we have incidental appendectomy in the elderly as the next item up for vote.

John?

MEMBER MORTON: Yes. So, this is another AHRQ measure, and it is looking at incidental appendectomy in the elderly.

This one generated a little bit more comment and questions than the previous one about perforated appie. It appears to be a relatively-rare event. The data that was submitted looked anywhere from about 20 to maybe mid-30s per 1,000 appendectomies. So, we are relatively rare. There was some supplemental data that was provided from the

State of Texas that was even lower, about 2 out of 1,000.

I think there is good science behind the reasoning for the measure. A lot of data has come out, particularly out of Mayo, looking at performing an appendectomy when it is not indicated leads to more problems than it is worth.

The one other question that came up in looking at the measure was there was an exclusion for colon cancer. I am wondering if there should be an exclusion for ovarian cancer since a lot GYN oncologists do try to remove the appendix at the same time.

What the measure has going in its favor is there are not that many elderly markers of quality out there. So, that is one thing.

I think the only question mark really from my end was whether or not this was something that was useful to measure, given that the numbers are fairly small.

CO-CHAIR TORCHIANA: Other comments?

MEMBER CIMA: I just had one followup. I know the exclusion for colon cancer, but there is a fair number of operations in colorectal surgery that you do that involve the right colon for polyps, benign neoplasms, colitis, that involve removal of the appendix as part of the procedure, not because it is abnormal. It is just stuck to the colon there and it needs to come out. And the pathologists often report it as a separate specimen. They will say appendix, you know, without diagnostic abnormality.

How would that be handled because colectomies are in the group that gets included? I mean I just was wondering about that.

MEMBER MORTON: Yes, I can address that. So, the exclusion list includes, basically, all colectomies that involve either

the entire colon or the right side of the colon. They also exclude pelvic exenterations. And there is separate exclusion for patients who have either a malignant neoplasm of the colon or malignant neoplasm of the retroperitoneal or peritoneal tissues or peritoneal carcinomatosis, basically, as well as secondary malignant neoplasms that involve those structures. So, I think those cases would be excluded.

The ovarian cancer might still get in here if the patient did not have a pelvic exenteration and if the patient did not have visible peritoneal carcinomatosis.

MEMBER CIMA: The only other comment I would make is there is a lot of data and interest in the perforated appie measure. Not a whole lot has been kind of done with the measure, at least in the literature regarding it.

So, the question is we are collecting them, but what we are going to do

with it?

MEMBER DILLON: So, in the sea of metrics that we are all forced to collect, is this one that we really need to do? I mean, is this really going to make a difference in terms of geriatric surgery or whatever?

MEMBER MORTON: Well, to Peter's point, I think this is one that might fall into that category about being a legacy measure, if you will. It is kind of, again, not a large volume. I am not sure if it really is impactful on the quality of care of the elderly.

MEMBER STAFFORD: The other question I had got back to the exclusions, and they are up there on the board. We talked about this is the conference call.

I am not sure why you would exclude a transverse colon resection, a left hemicolectomy, a signoidectomy, because those are operations where you wouldn't normally take out the appendix. It is not associated

with those types of resections. So, I was wondering why that was excluded.

CO-CHAIR TORCHIANA: Well, I am going to make a guess, and I am on thin ice here. But that the notion was that the downside of an incidental appendectomy is exposing the lumen of the large bowel, and that deed has already been done, I would guess, in what you cited.

MEMBER MORTON: That was the rationale, plus the possibility there might be some confusion in the coding as far as what particular segment of the colon had been removed.

CO-CHAIR TORCHIANA: So, Melinda, if I could ask, the importance question or the question of whether this remains relevant or whether this should become inactive or moved to the hall of fame? Do we address that when we do this first vote on importance to measure?

MS. MURPHY: Right. So, if the

group determines that it is not important to measure by virtue of the level of performance essentially being topped-out, so, then, it could be intended it from moving to an inactive status, assuming that the group also agrees that all of the other criteria, scientific acceptability, usability, feasibility, are met.

CO-CHAIR TORCHIANA: Okay. So, importance to measure: one, yes; two, no?

(Vote.)

Six, yes; fifteen, no.

So, let's pause at this moment, then. I guess the question is, does this imply that we are topped-out and it is no longer important to measure? Or does it imply that the measure itself is not of value and should be dropped?

MS. MURPHY: And that is the question that the group has to answer: have you voted that it is not important because it is topped-out? And if that is not the reason,

then we don't proceed.

MR. ROMANO: I could address the issue of topped-out. So, looking at the national data over a period from 1994 through 2007, the rate was 28.6 per 1,000 in 1994. In 2000, it was 26.7; in 2005, it was 20.5, and in 2007, it was 19.6. So, it has dropped from 28.6 to 19.6 per 1,000 over a 14-year period from 1994 to 2007, still about 2 percent overall.

CO-CHAIR TORCHIANA: So, I guess my impression, based on the discussion, was really not a sense that this had been sort of successfully put to bed, the topped-out hypothesis, but more that it was of dubious relevance and value, which would mean that we would not inactivate it, but, rather, we would eliminate it.

Is there a way that we can revise the vote to reflect those two alternatives?

MS. MURPHY: Well, we will capture the reason for the vote.

CO-CHAIR TORCHIANA: Okay.

MS. MURPHY: So, you do not need to revote. We just need to be crystal-clear that the reason for the no vote is other than being topped-out.

CO-CHAIR TORCHIANA: Okay. I was a no vote; it certainly was for me.

Does anyone feel otherwise that voted no?

(No response.)

Okay.

MS. MURPHY: So, we stop there.

CO-CHAIR TORCHIANA: Okay. Thanks.

So, now we have the first measure from the -- oh, I'm sorry -- from the ASC Quality Collaboration, hospital transfer/admission, 0265.

MEMBER WILHOIT: This measure assesses whether patients admitted to an ambulatory surgery center were transferred to a hospital or admitted to a hospital upon

discharge from the ASC.

The rationale for the measure is that about 80 percent surgeries are performed on an outpatient basis, and patients selected for ambulatory surgery are not anticipated to require hospital care upon discharge. Therefore, high rates may be an indicator that patient selection guidelines are in need of review.

Work Group A reviewed this measure, and it did have a number of concerns. We felt that the measure as written could potentially have unintended consequences: encouraging a discharge to home with instructions to go to the ER if there's problems, because if the patient is discharged home and, then, goes to an ER, that is a success from the standpoint of the measure; whereas, if the patient is directly admitted, that is a failure.

And so, we felt that a better numerator timeframe would be if the patient is

admitted to the hospital, instead of being at the time of discharge, if it were within some number of hours, maybe 24, maybe something shorter than that, but that the measure needed to include discharges within some timeframe or admissions to the hospital within some timeframe after discharge from the ASC.

Also, given the low admission rate in the data that were presented, the Work Group wondered if there is enough of a performance gap for the measure to be meaningful.

We also noted a number of apparent discrepancies in this submission, and these are all kind of fine-print, but play in.

The data were presented for a convenience sample of 526 surgery centers, but there is also comment that results are publicly-reported for 1,185. We weren't sure why we were only getting half the information, if it is being publicly-reported for twice as many as we got data on.

In the statistical analysis section, it didn't seem like the statistical analysis was valid. There was a comment of how much deviation it took from the norm to be an outlier, but it seems like that would vary depending on the population for a given ASC.

And, also, in some of the measures or in some of the data, the decimal point seemed to be misplaced and rates per thousand and percentages seemed to be confused.

So, there were a couple of fine points where the submission wasn't clear.

CO-CHAIR TORCHIANA: Now, then, David?

MR. SHAPIRO: Yes, we also have Susan White on the phone.

Susan, if you are on the phone -- I know you are on a short schedule -- did you want to speak especially to statistical issues that were raised?

MS. WHITE: Yes, I can.

Can you hear me okay?

MR. SHAPIRO: Yes. Please go ahead.

MS. WHITE: Okay. So, in a statistical issue, really the analysis that we were using was percentages that were rates by ASC. So, these are analyses of ASC and not the patient. So, that might be part of the confusion in the error rates.

I am sorry. The other timeframe questions -- I am sorry, I am boarding a plane; I apologize. The delay kind of threw me off.

All of the rates should be reported by percentages. I hope Donna can answer about the 500 versus the 1,000 for the public reporting.

MR. SHAPIRO: Thank you, Susan.

Donna, do you want to answer?

MS. SLOSBURG: I an not sure why there is that discrepancy. The only thing I can think of is that the 500 came from the ASC Association. And again, this is all self-

reported data. I could get that for you, but it is self-reported, and those 500 are reported in more detail than the 1100. Again, it is facility data that is aggregated, so we don't have specific demographics and specifics. So, that is where I think the 500 came from, is from the ASC Association's report.

MR. SHAPIRO: Yes, I want to apologize. That is a great point, and it is something that, unfortunately, we are not able to investigate here on the fly, but we certainly need to clear that up. And I apologize to you all for not having had that addressed prior to coming in here.

Let me just address one of the main excellent points that you brought up. That was the issue of the timing. And actually, this was alluded to with some of your other measures, and I was interested to hear that it still pertains to other sites of service.

And that is that, when we were constructing these measures, we were very concerned to only try to elicit data that was observable directly under control and reportable, therefore, during the continuum of care at the surgery center. So, as soon as that patient leaves the center, we have no idea whether that patient got into a car wreck or went home and, then, ingested whatever, got a fishbone stuck in their throat even though they had had a hangnail removal at the surgery center, in other words, something completely unrelated.

So, we felt that there were many, many, many factors beyond our control regarding the health and the eventual outcome of that patient's surgical procedure after they left the facility.

So, while I appreciate the notion of the unintended consequences, I also want to speak to the fact that this really has no pejorative implications in terms of why to

track this. This is really a global consideration for an ambulatory surgery center which, by CMS's definition, is meant to treat patients for a timespan of less than 24 hours.

So, it is very important that we have this data for our own centers. I think it is important to consumers, so that they know that the centers are doing what they are supposed to be doing. It is very important to the regulators, but much more importantly, it is to the facilities so that they can really keep a good eye on their patient practices. That covers pre-op, intra-op, and post-op.

And for pre-op, it is really patient selection. We want to make sure that we are taking the right patients into our facilities.

Intra-op, we want to make sure that we are not giving them too much medication so that they are unable to go home.

And post-op, we want to make sure that not only have we chosen the right

patients, but that we have arranged for appropriate postoperative care, treated postoperative nausea and vomiting and perhaps pain in a timely and appropriate fashion.

So, the issue of direct transfer is a lot because of the logistics of the surgery center, but we feel there was very, very much consideration, and we felt that the way that we did it, doing it only direct, rather than attaching a timeframe, was something that was worth tracking.

I will tell you, however, that Medicare does track both 24-hours and one-week return to surgery for these patients in the Medicare population. And we agree that those other kinds of indices, although they are specific for other issues, are also important to track, but they fall outside of the purview of this measure.

MEMBER SIPERSTEIN: We follow this at our own institution, and it is not one of these less-is-better metrics. I mean you

don't want a zero rate. It is like having a zero negative appendectomy rate.

But the way we use it is to track it longitudinally over time, where if there is a marked increase or decrease, then that really warrants a deep dive into individual patients to find out why there is a change in the pattern.

So, there may be a misunderstanding that less is better. I don't think that is the point of this metric.

MEMBER HALPERN: Is there a known acceptable rate of transfer?

MEMBER SIPERSTEIN: Again, it depends on the patient population, the type of surgery done in your particular ASC. I mean we have one ASC that does unilateral knee replacements. So, it all depends on the patient population that you are dealing with, but within your ASC, given a relatively-stable patient population, fluctuations in rate should trigger some alarms.

MEMBER CARPENTER: I would follow up with that and say, is this really a quality measure? The options are, do these patients at your surgery center with a slightly higher risk of admission or do them as an in-patient. And I am sure that their transfer from your surgery center to your hospital is an indicator that that is a poor quality or at least a poor outcome.

And, in addition, the rate is so low, it is .1 percent. Is that really something that is going to be valuable to measure differences between the surgery centers, if they are that low, or even differences over time?

It really has to do with the extent of things you do at your surgery center, how you are set up for these transfers, whether you do people with sleep apnea and things like that that have some trouble with postoperative blocks and things like that. Really, the options are doing

these patients at the in-patient facility, which may not be a higher-quality experience for them.

MR. SHAPIRO: Yes, I think that you all are raising very good points. And, actually, we talked about this last time we were here, when we initially presented this measure.

I still think that it is a very important measure for a facility to track. And that is what I was saying about it is not pejorative. It is really something that you want to trend and, then, do analysis on, and you certainly cannot do that deep-dive analysis or something along the lines of root cause to see what is it that is causing an increase or a decrease in your transfer rate.

We admit that they should be low, but within a facility, given a patient population and given the procedures that are performed in facilities -- remember, there is a wide range across ASCs -- this is something

that is very important, again, not only for patients and consumers, but something that the regulators are looking very closely at. And we have always been encouraged, especially by folks like CMS and at Medpac, to continue to track, and they use these statistics and look at those very closely.

So, I would encourage you to encourage us to continue to have this as an endorsed measure that we keep track of, again, knowing that a lot of times a transfer reflects actually good, and sometimes preventative, medicine, and it is always done in the best interest of the patient, But it is something that we want to keep a very close eye on, especially within a facility, for all the reasons that myself and others at the table have stated.

MEMBER HALPERN: So, how does it get publicly reported, then, to say that a given rate is not a bad thing? So, like let's say you are doing, I don't know, the

unilateral knees or people who are more sick. How do you report it in such a way, since this is for public reporting, that it doesn't become a punitive thing to that institution if they are doing the appropriate care?

MR. SHAPIRO: I think that what we have done in the past is we have the data reported on our site, and it includes an explanation such as some of the issues that we have raised at this table, to make it very clear to all that would go to the site, including our patients as consumers, what the implications are and, more precisely, what they aren't regarding any one of the indices that we track.

CO-CHAIR MORRIS: So, it sounds like we don't know what the appropriate target is for a transfer rate. And it also sounds like this isn't necessarily ready for primetime as a quality measure.

I understand that we are talking about it in terms of maintenance of a measure.

But if we don't know what the number should actually be, then, how can we say here's a target?

MEMBER MORTON: Can I get one point of clarification? Is this just for transfer from an ASC to the same hospital? To any hospital, right?

MR. SHAPIRO: ASCs are required by Medicare conditions for coverage to have a transfer agreement with a local hospital. So, that can be any one in the community.

MEMBER MORTON: They are not necessarily the same healthcare system?

MR. SHAPIRO: No, they do not have to be at all. Sometimes that is impossible. Of course, often that is the case, but that is not a requirement for regulation or for accreditation.

MEMBER MORTON: Thank you.

CO-CHAIR MORRIS: I absolutely agree with you when you talk about the importance of this measure. It sounds like it

is important to measure and track for your organization and for individual ASCS, but it is unclear to me how it is important to track as a quality measure. Can you explain that?

MR. SHAPIRO: Well, I am going to try.

I think it has a lot to do with the quality of care that we provide. And again, that starts with the preoperative selection and preparation of our patients. So, that includes making sure that they are someone that has someplace to go to for postoperative care. It also means that we have evaluated them correctly for comorbidities and that these are patients that can, indeed, be appropriately done for that procedure on that day in that facility.

And, then, it also goes to how we treat them intraoperatively, again, specifically, the anesthetic, but, also, make sure that there are no intraoperative complications as a result of the surgical

procedure itself.

And, then, finally, it really relates very closely back to the postoperative care in terms of how we are able to treat the postoperative pain, either with IV medication or pump, or both, and post-op nausea and vomiting, and any other thing that would have a patient be transferred to a hospital for treatment and/or evaluation.

CO-CHAIR MORRIS: So, it is a black box? It is a big black box that encompasses a ton of stuff, and we know zero is bad and 100 percent is bad. But it seems very fuzzy to me still.

Again, I completely agree with you that this is very important to understand and to dig into, but right now we are not creating a system for digging deeper. We are trying to identify clear, concise quality measures.

MEMBER HALPERN: I think what we are saying is, what parameters do you use? If this is a quality measure to be measured,

because that is what measure implies, how do you do it if you don't have a rate that is a target?

MEMBER CIMA: Just to follow up on that and what Richard said earlier, I mean you have a rate here that goes from zero to 2.3. The mean rate, as Jane pointed out, was .2 with a standard deviation of -- or .1 with a standard deviation of .2 percent.

You're doing a great job. And we would like to see it as zero, but we said earlier the STS was fine and it was topped-out at 95 percent. You guys are 99 percent doing the right job.

Why go through this exercise? I mean, for your institutions, you should track it, yes, and make sure, but I am not sure, because Arden's point is that you guys are doing a great job. And should we continue to make this something -- we could say it is something ASC should do every year, to check their rate. Are you going to put it on the

back shelf, or whatever we have been calling it, as something to follow? But what we did earlier, just sort of not retire it, but inactivate it.

But I think I am not sure what all the --

MEMBER MORTON: I would like to speak. I would like to speak up for the measure.

I think it is a good measure to take a closer look at because the consequences can be pretty severe. And it also looks at the patient selection. A lot of things go into play here.

It is not exactly germane to this discussion, but we have looked at ambulatory bariatric surgery. Yes, such a thing exists. And we have seen some very untoward results where a much, much higher mortality rate is associated with an ambulatory bariatric procedure.

So, I think this is something that

needs to get measured. There is not a lot of good data out there about ambulatory cases being done.

And I think the fact that we don't have a set rate for it, I don't think that is all that unusual sometimes in quality. Sometimes we are looking for those rates, and what is good now may not be good in the future. So, I would be careful about pegging it.

You know, I hate to keep bringing up bariatric surgery, but 10 years ago 1 percent mortality was acceptable. It is now .2.

So, it is something I think that bears merit to continue to look at because those cases are almost sentinel cases that give you some insight as to what is going on around patient selection or even the case itself.

MEMBER HALPERN: So, how about a delta for change? So, if you don't have a

specific rate, why don't you look for a specific rate of change?

MEMBER AFSAR-MANESH: I am wondering if, instead of even looking at rate, which we seem to have a little bit of a problem with, could we perhaps open up that black box by looking at best practices or guidelines that exist for various parts of this pathway? So, patient selection, what needs to happen intra-op, what needs to happen post-op.

Are there some evidence-based guidelines -- and I don't know; I'm asking -- is there something that we could use as processes to make sure that this is happening better as opposed to the rate?

MS. SLOSBURG: I think one of the troubles -- and I apologize; this is Donna talking -- is that there is not a lot of evidence out there for ambulatory surgery, No. 1.

No. 2, we just don't have a lot of

data.

And, No. 3, again, these are 1200 facilities out of 5200. Again, this is voluntary. When you talk about topped-out, these are centers that have been doing this for multiple quarters.

I think that if you looked at the rest of the population, maybe their rates are much higher, but we don't know that. And again, I think from a quality standpoint, if a center has a high transfer rate -- and I know you all asked the question, what's high? Personally, I think anything over 2 percent. You definitely have to do some digging in and finding out what is going on in that center because maybe you are doing a lot of higher-level cases, bariatrics or total knees or total joints, but, also, maybe you are just not doing a really good job on your pre-assessment. That says a lot to the care that is going on in those facilities.

I don't know how to get the point

across, but I am telling you all for our industry this is a huge measure, and it is really needed out there.

MEMBER HALPERN: I think we all agree that it is important.

MS. SLOSBURG: It is just the number.

MEMBER HALPERN: Yes. So, how about something including, since it seems like the people who are involved in doing it now have done a good job, so maybe the target should be to get more involved, more centers involved.

MS. SLOSBURG: And that is actually our goal. We are working on a registry this year -- we do have some funds -- to allow anybody to come in and put their data in, so we can get more centers, because we just don't know what is going to happen with CMS. We don't know if it will be January 2012, 2013. We don't know.

CO-CHAIR TORCHIANA: Could I butt

in for a second? I would like to recognize speakers because there's a few people who are trying to talk who aren't as quick on the button.

(Laughter.)

DR. BURSTIN: Just two overarching issues, I think in some instances this is very analogous to some of the safety indicators. A low rate doesn't necessarily mean it is the thing you shouldn't track.

We also do have some good examples of things like episiotomy, not a good thing, but we don't actually know what the right rate is. It is kind of dependent, but it is still an important indicator that you might want to track over time.

My third issue that I wanted to raise, again, a cross-cutting issue, is I think the idea of having a measure stop at the time they leave your door has sort of changed and evolved. There is a definitely shift towards shared accountability. I suspect the

number is significantly higher, and I would be curious to hear what the research would suggest, if you actually looked beyond that discharge time.

Certainly, many of us are held to 30-day accountability and things like that. So, the idea that as soon as they leave your door, you are kind of done I think is kind of old thinking.

I think it is really important to think about, whether it is a seven-day window, or whatever the case may be, if you overmedicate somebody and they are still woozy and they fall in your parking lot, sorry, but I do think that is the responsibility of the ASC.

MEMBER SIPERSTEIN: But I think, in aggregate, the measure makes sense, you know, in that it is forcing the ASCs to pay attention to this. And if the whole goal of the metric is to try to improve quality, it is at least focusing efforts because a rate,

again, that is too high or too low may prompt you to do a root-cause analysis.

I think some of the scientific issues are that we are trying to dissect it apart a little too much and go be their quality committee and do their root-cause analysis for them.

I do see that if you have got a group that has got a 5 percent rate, yes, that is going to raise a red flag on a comparative basis. But I think, in aggregate, it is not a perfect metric, but at least it is encouraging the ASCs to focus their attention on the area that does have some important potential quality impact.

MEMBER FINDLAY: If you were to publicly report this by institution, would consumers and prospective patients be able to make sense of and discriminate between centers, based on the data at this point?

MR. SHAPIRO: I think that is an excellent question. That is why, as I said,

it is publicly reported now. It is on the ASC Quality Collaboration website as an aggregate number, not by facility.

MEMBER FINDLAY: But not by --

MR. SHAPIRO: Right.

MEMBER FINDLAY: If it were to be reported by facility, as maybe you guys hope in the future, would it be meaningful?

MR. SHAPIRO: I think it would be very meaningful, but I think an informed consumer would very, very legitimately ask the exact same questions you are. And they are great questions to ask, but I just maintain that that does not invalidate --

MEMBER FINDLAY: No.

MR. SHAPIRO: -- the reason that we should keep track of this.

MEMBER FINDLAY: I am not suggesting it does. I just wanted that piece of information.

MR. SHAPIRO: Yes. No, it is a great question.

CO-CHAIR TORCHIANA: I think Robert was next.

MEMBER CIMA: I mean I spoke earlier about my feelings about this, but to follow up on Allan's point, it is that, you know, if you put someone out the door and they end up in the emergency room that night because they are having pain, you haven't done an adequate job of preparing them for their pain control.

In our practice, as a colorectal surgeon, 15 percent of my male patients who have anal surgery end up coming back within 24 hours because of urinary retention.

I mean, you know, you are doing a procedure. I am held responsible at the main hospital for readmission rates; everyone else who does hospital-based work is held responsible for readmission rates.

All right. So, let's say there is a need to everyone to report it. Then, it just can't be walking out the door is the end

of the road. If they show up in a hospital -- I mean I get in trouble for 30-day readmission. I get 30-day deaths. I have had two patients in the last five years that have been killed within 30 days in car accidents. It still comes back to me.

So, I mean, like Allan has been saying, you know, responsibility does not end at the end of the procedure or when they walk out the door.

MEMBER CARPENTER: I was just going to get back to I just think this is being reported as a quality metric, and I really don't think it is a measure of quality for the individual patient.

We know that most surgical outcomes have high quality in surgery centers. So, in general, the quality of surgery is very strong there.

If you have an option of in-hospital or at a surgery center, I can lower this rate by doing everybody that is a

slightly higher risk in the hospital. It will drive up costs. It will lower this number for me. But it won't provide better quality for my patients. And my patients, if they are comparing two centers, won't be able to tell really what that means for them.

So, I do think it is important, and I do think it needs to be tracked. I just don't think it is a quality metric. I don't think you have presented any evidence that those patients that are transferred back have worse outcomes than if they had been done in a hospital, which is really the question.

MEMBER DUTTON: I think this is a very important quality measure. I'm sorry, Jim, but from the anesthesiologist's perspective, you have to collect and examine these people. You have to know what your rate is over time. Seventy-five percent of the anesthesia we do now is in the outpatients, and knowing that they stay outpatients is absolutely critical. It is pain management.

It is management of nausea and vomiting. It is management of some surgical complications.

I do absolutely agree with Helen, though, you have to push it out to 24-hour, all-cause admission. You are calling all those people back the next day to get your satisfaction data. You can find out whether they got admitted or not.

CO-CHAIR TORCHIANA: So, I am going to take the prerogative of trying to summarize where we stand, and, then, trying to figure out if there is a way to resolve.

I guess a couple of things that I would observe is that this is one of those things that calls the question of what does NQF certification actually mean. Because there is no reason why the ASCs could not continue to measure this measure ad infinitum without NQF endorsement. So, what does the NQF endorsement actually do?

I would guess that maybe the NQF endorsement in this case helps to get the 1200

up closer to the 5,000, which would probably be a good thing.

I can't think of a downside of it, unless we feel like we are endorsing a truly inferior quality measure. And it seems like there are strong opinions on both sides of that topic.

And so, I think my opinion of this would be we ought to go ahead and vote, see what the vote turns up, but I hope that the folks in the ambulatory quality world in these centers take this advice.

And particularly, I would say the early argument, David, that you made that, you know, if somebody goes home and they cut themselves in the kitchen, or some other completely unrelated event, I think that one is a little bit specious because that exists really for every quality measure that we have once somebody is discharged. There is always a baseline rate of mayhem in the population once they get out of our doors. And so, that

is a problem with every quality measure that you try to follow up, once a patient is discharged.

So, I would urge -- and it may be through the satisfaction survey method -- that you take to heart the advice that this would be a much more substantive measure if you actually had a 24-hour window at least around the admission piece. And, hopefully, if we vote this in, next time you come back you will have done some work on trying to accomplish that.

MEMBER WILHOIT: You raised the point of, what does NQF endorsement add? Well, the NQF endorsement, according to the criteria we are using, has to do with internal quality improvement and, also, with public reporting.

I think it sounds from the discussion like, you know, there is real agreement, I think, that this or some variation is very relevant for internal

quality improvement. I think where the issue is is the public reporting, and that is what the NQF endorsement adds, is a blessing on public reporting. And it just doesn't seem like it is ready for that.

MEMBER ROGERS: I would like to echo, David, what you said about the imprint of NQF. What you do have is very nice data and very acceptable data on the centers who, in fact, have reported, but you have no knowledge of what goes on in the other ones.

I think that one of the things that this Committee might want to consider in its charge is exactly that issue: could the approval of a measure which may not be perfect, but it is a good measure, because I heard that, give them the kind of a little bit of extra muscle to get more people in? I think that that is a very important responsibility for this Committee and NQF in general. So, I would support it on that basis.

CO-CHAIR TORCHIANA: A final comment from our developers. Then, I think we are going to vote.

MR. SHAPIRO: Yes, I would just remind this group that we take very much all the issues that you have raised. I think they are excellent, and they are actually ones that we have discussed during measure development.

But I can tell you from being, I guess, a pioneer in the industry of ASC quality reporting, and we are at the very, very embryonic stages, that the last three years of having the endorsement of this body has really allowed us to get great mileage. We started at zero three years ago, and we are now up to 1600 centers and very close to getting even more than that, as we are working on incorporation in a registry and an AHRQ and other things.

So, I can't tell you how much it has meant to us. Even though this may be an imperfect measure, as maybe our others are,

but to have had the approval and the endorsement of this body has been an exceptional boost for us and for our patients. So, I appreciate that.

CO-CHAIR TORCHIANA: Thanks.

Let's vote on importance. One is yes; two is no.

(Vote.)

We have at least one person out of the room. I think 20 is the number.

So, fifteen, yes; five, no.

Scientific acceptability, a scale of 1 to 4.

(Vote.)

So, this one is a bit more trying. Two, completely; ten, partially; six, minimally; two, not at all.

So, usability.

(Vote.)

Six, completely; nine, partly; three, minimally; two, not at all.

Feasibility.

(Vote.)

Thirteen, completely; seven, partially.

Does the measure meet all the NQF criteria for endorsement? Here is a yes, no, or abstain.

(Vote.)

Thirteen, yes; seven, no.

MS. MURPHY: May I ask a question? At this point, as you did with measures in Phase I, do you want to ask that the developer come back to you with information about the 24-hour period?

(Chorus of yeses.)

MEMBER WILHOIT: And could we also ask that the measure take a look at the statistical analysis and the decimal points? There's rates per thousand and percentages that are mixed up, and the statistical analysis I think is lacking in a major way.

MS. SLOSBURG: For clarification, can I just ask, are you asking us, then, to

leave the measure as is and move it out to 24 hours? Is that what I am hearing?

CO-CHAIR TORCHIANA: I think we would like to hear a plan as to how it could be moved out to 24 hours, but for the time being the measure is endorsed.

MS. SLOSBURG: Okay. We will do that.

MS. MURPHY: Let me clarify. What you voted on is whether or not the measure meets criteria for endorsement. The feedback that you bring back to the group will occur, and they will consider that prior to making a recommendation for endorsement.

CO-CHAIR TORCHIANA: For our next measure, 1517, statin therapy at discharge.

MEMBER WILHOIT: This measure assesses the percentage of patients undergoing infrainguinal lower extremity bypass who are prescribed a statin at discharge.

It should be noted that the measure is defined in terms of medical record

data. There is no clear indication in the documentation of how frequently a registry is used. Without a registry, the data would not be available electronically. So, the Work Group that reviewed this had some questions about the feasibility of implementation. There was no mention of the number of participants currently in the two registries that exist.

One other question, a couple of other questions that came up in our discussion. One is that the measure is based upon a guideline. However, the guideline that is cited recommends statin use based upon the level of LDL control, and the measure is measuring statin use regardless of the LDL level.

Also, the numerator and denominator timeframes lack precision.

CO-CHAIR TORCHIANA: Comments from the developers?

MR. KRESOWIK: This is Tim. I am

still on.

I don't know if Lindsey can comment further.

This measure, it is coming out of the Northern New England registry, and they have a number of sites. I don't know the exact number of participants.

I guess the only other thing I could comment on with reference to the guideline is, I mean, that is the existing guideline, but certainly emerging evidence is suggesting that LDL level is not as important in terms of determining the benefits of statin. As you all know, a lot of times the guidelines lag a little bit behind the current state of the evidence. But that is certainly, I think the use of statin without reference to LDL is certainly the current trend and I think will be supported in the future guidelines.

CO-CHAIR TORCHIANA: Other comments?

(No response.)

Should we be prepared to vote, then?

Carol, do you have any other?

MEMBER WILHOIT: No.

CO-CHAIR TORCHIANA: Okay. So, if we could vote on importance to measure and report?

(Vote.)

It looks like we are not going to get past 19 here. Nineteen, yes; one, no.

Scientific acceptability, a scale of 1 to 4.

(Vote.)

Eight, completely; eleven, partially; one, minimally.

Usability.

(Vote.)

Fourteen, completely; five, partially; one, minimally.

Feasibility.

(Vote.)

Thirteen, completely; seven,

partially.

And does the measure meet the NQF criteria?

(Vote.)

Nineteen, yes; one, abstain.

I am going to hand the rest of the measures to Arden. I am not going to hand the rest of the measures to Arden.

(Laughter.)

All right. So, the next one is AHRQ, abdominal aortic aneurysm volume, 0357.

MEMBER SAIGAL: Okay. So, that is me.

So, these are a composite pair with the one that is following. Should I do them together?

CO-CHAIR TORCHIANA: Please.

MEMBER SAIGAL: Okay. So, the first measure is a volume measure of AAA repair based on hospital discharge at the provider level. And the companion measure is discharge in which a AAA repair was done,

endovascular and open repairs, in which an in-hospital death occurred.

I think that, in terms of the first measure, in terms of importance, this document, there is a lot of cutting and pasting going on, and it was a little hard to interpret it in certain places.

In terms of looking at the performance gap, they report the numbers of procedures done by age and race, but not by provider. So, I am not sure if the person doing this format, what they were trying to get at with that data.

But we know this is a very important measure. One of the earliest measures for quality is AAA repair volume.

They didn't have much in terms of disparities, either. They lump in endovascular and open repairs in this. I am not sure. There are some comments from our group that those could be reported separately because of the trend towards open repairs

happening in more complex cases now. But there may be some dipstick value to the overall discharge number that they don't really talk about that much.

They have, in terms of usability, in this measure they talk about a report that was generated. There is a lot of cutting and pasting of methods of the report, but there is no data as to what the report said. So, I would like to hear more about that.

And, then, in terms of the mortality measure, there were two risk stratification methods I read about. One was in 2A.14, which was what they used in the NIS analysis that looked at gender and age, I believe, as the only stratifying variables for mortality, probably because that is all that is available in HCCUP.

And, then, the other one was a CMS model that they referred to that was more detailed in terms of a model. So, I wasn't clear as to what they were proposing be done

for risk stratification for mortality.

And, then, the last thing that was kind of troubling to me in the mortality measure was that in Section 2B.3 they go through a long thing about how they defined signal-to-noise ratio in their measures. And they mentioned that the mortality measure at the provider level had a low signal-to-noise ratio. So, I wasn't clear if they were -- it is a kind of sentence fragment as well. Maybe it was some typos in there. So, it was really hard to interpret that, but it raised some alarms in my mind as well.

MEMBER HALPERN: I also just wanted to add about the endovascular versus open repair, and lumping them together, because open repairs are not what they were 10 years ago. And a lot of the literature that is quoted is from 10 years ago.

Open repairs have become more complicated because the vast majority are being done endovascular, and when they are not

being done endovascular, it is generally because it is a short neck. It may mean a superrenal clamp. Or it is a bilateral iliac artery aneurysms that are not well amenable to endovascular repair. So, they are much more complicated cases than they were in the past.

And, also, in terms of volume, there is actually a paper that just came out last year in Circulation that looked at volume in relation to outcomes. And actually, in endovasculars, the volume threshold is 10, 10 cases per year where they started to see a difference, more or less than 10.

So, the volume, and it is much more dramatic in open repairs, again, I think because open repairs have become so much more complex. So, I think lumping them together is not a good thing.

MEMBER MORTON: Do we know what the breakdown is in the vascular versus open?

MEMBER HALPERN: I would say probably now at least 80 percent are done

endovascular, at least. I am not sure if the SVS tracks that. They could probably answer that better. But I would say I know in my practice it is probably 80 percent, and those that aren't, you know, again, the more complicated patients anatomically.

MEMBER MORTON: Renae mentioned about emergencies. Almost all of those are being done open or?

MEMBER HALPERN: It depends on the center. And it may be a reason why to take out the ruptures. Those centers that have a process in place to do them endovascularly do them, tend to do them endovascularly.

And when I say a process, you have to have grafts onsite. You can't be waiting for a ref to come and bring you grafts. So, if you don't have stock in grafts, you may not be able to do them endovascularly unless it is a stable ruptured aneurysm, you know, somebody who presents with abdominal pain and has a rupture on TT, but is not unstable. But for

unstable patients, unless you are set up to do them, you are not going to be doing them.

And, then, of course, the open ruptures are much more disastrous than the endovascular ruptures. So, again, another reason to separate them out.

MR. KRESOWIK: This is Tim Kresowik. I am not here to comment on these measures, but I will.

As I discussed in the introduction to the other measures, the SVS and the people involved with quality from the SVS would not support either of these measures based on the scientific validity.

Certainly, the risk-adjustment methodology based on either NIS or administrative claims data is far from valid. I have already alluded to the previous problems with volume, and there has been multiple studies that have shown that, certainly, when you try to come up with a clear threshold, it is very problematic. And,

then, the trend is there.

But once you start risk adjusting, you lose a lot of the validity, if you will, and it always adds a perverse incentive, as I said, because people tend to try to meet these volume thresholds. And the only way you meet these volume thresholds is by operating on more and more patients who would be better off not having a procedure altogether.

So, I just would register that, from the Society for Vascular Surgery point of view, we would not support these measures.

CO-CHAIR TORCHIANA: Could we hear from AHRQ?

MR. ROMANO: Yes. So, we have actually done extensive analyses here. There was discussion of these points as well during the original endorsement process several years ago.

And there are actually strong methodologic reasons for the choices that were made here. So, in terms of, for example, the

increased use of endovascular procedures, the national data, of course, lag several years behind leaders' practice. The last national data that we looked at showed about 60 percent endovascular, but the trend was still upward. So, it wouldn't surprise me if that was up to 75 percent now.

But what we have seen is that, as people have converted to the endovascular approach, the short-term mortality has decreased. Now there is some argument in literature about whether endo leaks and other long-term complications may erode some of that early benefit.

But the fact is that, from the patient-centered perspective, that it is important to combine the two types of procedures because, otherwise, you miss the fact that there has been a temporal trend towards decreasing short-term mortality, which is largely attributable to the switchover from the open approach to the endovascular

approach.

So, this is a specific process that has been implemented within most hospitals that has improved short-term outcomes. So, when we see that, we want to give hospitals credit for that improvement. And the way to do that is by constructing the indicator in a patient-centered way, so that it reflects the patients who are coming into the hospital for their AAA repair.

In terms of the issue of the ruptured and unruptured, so this is also, of course, a very important issue. I would say that the risk-adjustment model, and if you are interested, I can show a copy of it, but it certainly does include as one of the most important factors the rupture of the aneurysm. So, that is taken into consideration in the risk model.

Again, putting these two types of procedures together is based on three premises. One is empirical literature showing

that surgeons' and hospitals' experience with ruptured cases improves outcomes for all cases, and vice versa.

So, generally, it is the same surgeons who are operating on ruptured and unruptured cases. And the experience that they accumulate, basically, carries over to improved outcomes for both ruptured and unruptured cases.

So, by combining the two together, the performance of the indicator substantially improves in terms of the reliability and the ability to discriminate amongst providers. Roughly 50 percent of the variation that is linked to the volume outcome association, roughly 50 percent of that is attributable to ruptured aneurysms; roughly 50 percent is attributable to unruptured aneurysms.

So, we see this tracking together, that the centers that have better outcomes with ruptured aneurysms also tend to have better outcomes with unruptured aneurysms.

So, again, this is sort of a methodologic rationale for combining the two and dealing with the obvious difference in risk through adjustment in the statistical model.

MEMBER HALPERN: How much is ruptured weighted in your model?

MR. ROMANO: So, the ruptured cases constitute roughly 10 percent of the cases and roughly half the deaths.

MEMBER SAIGAL: Could I ask, could you comment on the notations you guys have in there about the signal quality of the mortality measure? And also, if the model that you published in the application is not the correct one, it would be nice to see the one that you actually are going to use.

MR. ROMANO: Sorry. The model is the one that is currently in use, and I am pulling that up momentarily, but my internet is a little slow here.

And I am sorry, I didn't prepare this myself, so I will look over it and have

to respond in a minute.

MEMBER CIMA: If the purpose of these measures is to inform the public as well as the practices, most people don't choose their hospital for a ruptured aneurysm.

(Laughter.)

So, although I understand the methodology. You know, it is somewhat like a trauma center. You know, you do a lot of ruptured aneurysms; you get good at doing that.

But to inform the public about where to get your aneurysm repair, is it necessary to have that in there? I mean we have tried to segregate in other measures acute CABG. You know, salvage CABGs and thing like that have been separated out. Can't that be done here, too? I mean I don't understand why it would -- I understand your methodology about tracking centers, but I don't understand why it has to be done in this measure.

MEMBER HALPERN: And again, I

point out there is in terms of just using volume, also, there is literature now showing -- and this is from Circulation 2010 by Landon, et al -- and they showed not really a significant difference over 10 cases for endovascular elective repairs versus open. And again, because open repairs, open elective repairs are becoming more complex, and like Dr. Cima said, you know, you don't choose the hospital where you wind up with a rupture.

MEMBER MORTON: Just out of clarification, for CABG, do we segregate it this way between emergency and elective?

CO-CHAIR TORCHIANA: It is in the model. There is also a lot of controversy around whether or not patients who are receiving CPR, either on their way into the OR or immediately prior to going to the OR, should be included. But, as of right now, they are included in the model.

MEMBER CIMA: There are some exceptions in the sense a salvage operation

doesn't need a mammary. So, they have taken those patients out, you know, those high-risk, emergent-type patients, those are out of the pool. I can't think of anything more emergent than a person with a pre-ruptured AAA.

MEMBER SAIGAL: Can I add, so this is a provider-level measure, according to the AHRQ application? And they state that the hospital-level measurement is unreliable, although the usability report they reference talks about informing policymakers about hospital quality. So, that is also a little confusing.

MR. ROMANO: I am sorry, I am having a little trouble finding the specific section of the document that you are referring to, if you could clarify that? And I don't know if --

MEMBER SAIGAL: 2B3.

MR. ROMANO: 2B3?

I don't know if the statistician on our team is on the call. Jeff Geppert?

(No response.)

I guess not. Okay.

In the meantime, I am going to pull up the risk-adjustment model for AAA mortality. So, that model includes gender, age categories starting at 65 up to 85 and over. It includes ruptured versus unruptured aneurysms. It includes major diagnostic categories that reflect comorbidities, major comorbidities.

And it is based on the APR-DRG risk-adjustment scheme. So, it includes three APR-DRGs that basically correspond to features of the aneurysm and associated cardiovascular conditions.

We are looking at 2B.3 here. So, the signal ratio is the proportion of the total variation across providers. It is truly related to systematic differences in provider performance, and it is 31 percent, which is lower than some, but higher than others. So, it was high enough to meet our threshold for

inclusion, but some other indicators are in the range of 50 percent or over. So, this is a policy debate as far as whether 30 percent signal is high enough.

MEMBER CIMA: What was the value for the weight of the rupture in your model? I mean you put in, you cited off about a dozen different things. But the main issue here is ruptured should be included or not.

MR. ROMANO: Right. The point estimate in a logistic model with a C statistic of 0.909 is 1.8. So, that would translate into an odds ratio of e to 1.8 or, roughly, 5, I am thinking. So, I think that corresponds, roughly, to the absolute difference of about 5 percent versus 30 percent or so. So, yes.

And the overall C statistic of .909 reflects the discrimination of the model, and that is generally considered very good for risk-adjustment models based on these type of data.

MEMBER MORTON: Actually, I like the argument about, if you are good at doing some open cases or if you are good at doing endovascular cases, you might be good at doing emergency cases. So, I think Dr. Romano's point there, actually, for me, had some traction, that it is good to include them.

I understand that they are all kind of different, but the fact that we don't know where these cases are coming from and what those providers do, I think it sounds like it might be useful to keep them in there, the emergency cases and the open cases.

MEMBER HALPERN: I am not saying we shouldn't look at open, but I think there is a vast difference now between endovascular and open cases. And, remember, this is publicly reported. So, if I were a patient, I would actually rather know, if I had a complex aneurysm, I would rather know about the volume of open than the total volume because maybe somebody dose 100 endovasculars

and two opens.

MEMBER MORTON: Because what I am hearing is part of this --

MEMBER SEARS: This is Nick Sears.

What about the conversion to open from an endovascular? Are we considering that at all as well?

MEMBER HALPERN: No, I don't think that is considered in actually any of the measures.

MEMBER SEARS: Yes.

MEMBER HALPERN: I think the rate is very low now. I remember seeing a paper about it not too long ago, that the rate of conversion is under 1 percent. I think our selection criteria has gotten better.

MEMBER SIPERSTEIN: But I think what I am hearing in this discussion is that there is an overall mortality rate for these vascular procedures, and you almost have a 2x2 matrix of open versus endovascular and elective versus ruptured.

And looking at the overall and looking at the subsets, I think both have different values I mean, if you are looking within a hospital system in terms of are you appropriately giving the right therapy to the right group of patients, you want to look at your overall numbers. Again, if you are a given patient who is faced with having an open versus an endovascular, yes, you may be interested in looking at the details.

So, I think maybe what we are doing is we are interested in the overall number, but realize some of the inherent weaknesses, and, then, maybe the stratification into those subcategories would be helpful to dissect this apart.

MR. GEPPERT: Patrick, this is Jeff Geppert. Can you hear me okay?

MR. ROMANO: Yes, now we do. Thank you.

MR. GEPPERT: Yes.

MR. ROMANO: Sorry.

MR. GEPPERT: Just on that last point, so part of this process, starting from when the indicator was originally endorsed, there was a series of communications with folks at SVS. And one of the things that AHRQ said they would do in response to the points that are being discussed is allow for the stratification of the measure in the software based on the open and the endovascular and the ruptured and the unruptured. So, this kind of 2x2 stratification that you just mentioned is a component of the current software.

CO-CHAIR TORCHIANA: So, of course, the problem with trying to stratify and have more homogeneous populations for comparison is that the number in your cells goes down and the ability to discriminate also goes down.

This is not a measure generally, nor are the pancreatic, nor are the esophageal measures, that have an enormous power to discriminate providers in the first place.

And so, this is one of the challenges around the desire to have good outcomes measures for these high-risk or relatively high-risk interventions on the behalf of the public, is that the aspiration for what the public wants versus what is actually realistically deliverable, there is kind of a big gulf there.

So, on this one, I think I am going to look to Melinda again. I think we should take this as a vote on importance or measure or we should send it back for more refinement. And I think that is an important discussion.

MEMBER HALPERN: I think we were, also, at some point talking about harmonizing the volume and the mortality into one.

MS. MURPHY: And discussion of harmonization with related measures comes after the vote related to meeting the criteria, but before a vote on recommendation for endorsement.

MR. ROMANO: And in this case, I would say that this volume measure and mortality measure do have the same definitions. So, at least these two measures are harmonized with each other.

And in this case, we have done extensive analyses of these questions, of looking at open versus endovascular, and so forth. And I think the Chair has summarized the issue very well, that we feel that the loss of reliability, the loss of discrimination in separating out half the deaths is not justified empirically. So, that is our view.

MR. KRESOWIK: This is Tim Kresowik.

I was going to comment on this endovascular versus open. I think it has to be viewed in terms of where we are today versus where we were when these things were first being looked at, in that there was, if you will, perhaps an option that people could

go one way or the other.

Then, I would have to say that in 2011 virtually nobody, if you will, is getting an open repair if they are a reasonable endovascular candidate. And so, I mean I think these procedures really have become quite distinct. They are not the same population. The patients that are getting open repair in 2011 are ones that have anatomic configuration that does not allow an endovascular repair.

So, I mean I think the environment has completely changed. And so, I think that has got to be taken into consideration as this is viewed.

MEMBER SAIGAL: A technical question. These things are composite measures. So, do you have to vote the same way for both of them?

MS. MURPHY: No. They are recommended for use as a pair. They are endorsed individually.

MEMBER SAIGAL: Okay. It is concerning that 70 percent of the variance in the provider level is noise in the measure. So, if you are going to use this for reporting about providers, that is a lot of noise, I think.

MR. KRESOWIK: Although that is similar to a lot of measures that are publicly reported, including some of the ones that we have talked about today for pancreatic and esophageal.

The other thing that may be a comment that is more appropriate for the public commentary, but in order for endorsement to be a meaningful process, it has to be something on which people can rely for planning and for decisionmaking. And this has been a long, three-year process from when the measures were originally endorsed to the point now where they are scheduled to be included in Hospital Compare.

And there have been publications

and notices in The Federal Register, you know, public comments to The Federal Register notices, several rounds of hospital reports that have gone out for hospital comments. This has been a several-year process.

And so, to change a decision based on a lot of -- you know, I understood the last point about there have been trends in the last couple of years that maybe merit reconsideration, but to make a decision based on a lot of factors that have been previously considered, were discussed with people like SVS, you know, in some sense, it really pulls the rug out of a process that the federal government has spent a lot of money and time and effort on. And so, it is just something that should weigh very heavily, in my opinion.

MEMBER CIMA: So, what exactly are we doing here, then?

CO-CHAIR TORCHIANA: The body language around the table did not receive that one well, since you are on the phone and can't

see it. I think justifying the continuance of a bad measure because a lot has been invested in it is not a very compelling argument.

MR. KRESOWIK: That is why I said it was probably more appropriate for the general comment period.

MEMBER MORTON: I kind of liked Allan's idea about looking at this on overall, all the individual ones. Because really what you are talking about here is the care of the patient with an abdominal aortic aneurysm, whether it be open, where you have some anatomy that drives that decision, or it is EVAR, or if it is emergent. I think all of them should be included.

You know, you have the overall because these generally don't occur in isolation. You just don't do opens and you don't just do EVARs and you just don't do emergents, right?

MEMBER HALPERN: I understand that, but to report them as one lump I think

does disservice to the differences between them. And if you are going to make judgments based on lumping them all together, you are not going to get a correct judgment.

CO-CHAIR TORCHIANA: So, I think we have covered the pros and cons here pretty well. And I guess I think the logical place to go is to a vote unless somebody asserts otherwise.

(No response.)

Okay. So, let's first answer, does the measure meet NQF criteria for importance to measure and report?

(Vote.)

MEMBER SEARS: Are you voting on 0357?

CO-CHAIR TORCHIANA: Yes, 0357, as we talked about it.

It is a dead heat, 10 to 10.

MEMBER SEARS: Do you want me to break it? Nick Sears.

CO-CHAIR TORCHIANA: Okay. Well,

let's continue through the remaining portions of the vote.

Ah, Dr. Sears, can you break the tie?

MEMBER SEARS: Yes. I would say no.

CO-CHAIR TORCHIANA: Okay. Eileen is also on the call?

MEMBER KENNEDY: I am.

CO-CHAIR TORCHIANA: Would you care to vote yes or no?

MEMBER KENNEDY: I am actually ask to abstain.

CO-CHAIR TORCHIANA: You're going to abstain? Okay.

So, do we proceed with the remainder of the vote, Melinda? Too tight?

DR. BURSTIN: It is so clearly a split, though, and it is hard to argue there is a consensus on that first vote. So, I think either have more discussion or see if there is additional information you want to

get from AHRQ and discuss on a subsequent call. But to just table it at this point seems a little, it doesn't sit right.

MS. MURPHY: So, some of the discussion was around the question of having it split and having open versus endovascular.

So, are there some conditions or requests that the group has of AHRQ to bring information back to reconsider it?

MR. ROMANO: And if I might say it, in the context that the current measure and software does support stratification, so that is an option that is available to users.

What could be done that is not currently done is separate risk-adjustment models could be developed for the different cells in that 2x2 table.

MEMBER CIMA: Well, I mean that is one option, but I mean I think the general consensus is people are viewing these as two very distinct procedures now. And if you are doing EVAR, you should have a mortality rate

of "X", and if you are doing opens, you should have a mortality rate of "Y".

MEMBER SEARS: And if you want to tie volume to them, you could tie volume to them. But I am not big fan of volume these days.

MEMBER CIMA: No, I am just saying, could that be, rather than stratifying, just have them separate? Because I know in our health system all the open ones come to our institution. All the surrounding hospitals no longer do them. So, it is sort of strange, you know. And all the ruptures come to us. So, we get the worst of both worlds because no one touches them. And so, that is the reality.

And I think we are comparing apples and oranges. As a surgeon in the room, I view them as totally different patient populations and totally different procedures.

MEMBER HALPERN: And as a vascular surgeon, I concur with that.

MEMBER WILHOIT: This measure is the volume measure, and if 75 percent or so of the cases are endovascular, the numbers are so small on volume anyway, if you separate out, it seems like for the open procedures there's not even enough left to report on volume.

So, in terms of this measure, 0357, which is volume, does it make sense to make it just endovascular?

MEMBER HALPERN: No. Again, if you look at the last paper that came out from Circulation, the volume effect on mortality for endovascular is very small. The threshold where they saw the biggest difference was more or less than 10, 10 surgeries a year. So, the volume effect is very small.

The volume effect is much higher for open. So, I think it would actually, even though opens happen less frequently, they have become so much more complex that I think it is still reasonable to look at that volume. That is where you really see the differences in

volume. At each quartile, the mortality goes down with the number of volume that you do.

MEMBER WILHOIT: But looking at the data in the volume measure, those 10 cases -- and this is combining open and endovascular -- 10 cases comes between the second and third quartile. So, you have an awful lot of facilities doing less than 10, no matter how you count it.

The second quartile is 5.6. The third quartile is 13.8. It is in 2F3.

CO-CHAIR TORCHIANA: So, I think it is worth pointing out that this is, in fact, just a volume vote, and we also need to look at the mortality rate vote.

But I guess I would absorb from the discussion that a reasonable charge for the developers would be to come back with a model that separates endovascular from open for both volume and for repair, in the absence of another suggestion.

MEMBER HALPERN: I will also

mention that there are papers out there -- actually, one written recently by my colleague down in Tucson, Joe Mills -- where their mortality rates, although they didn't hit thresholds, were equal to the ones at higher-volume hospitals.

So, at one point -- we have discussed this with some of the other cases -- it is the process rather than just the volume.

MR. ROMANO: So, we are certainly able to report back on the implications of separating the volume according to open versus endovascular and reporting those separately.

CO-CHAIR TORCHIANA: And I would suggest, also, the mortality, we could go through the exercise of voting on mortality as well.

So, should we do that, Melinda? Do you suggest we do? It is the same issue.

MS. MURPHY: Right. So, I would say no, and we would get the information based on what you just requested and, then,

reconsider the two of them at that time.

MR. ROMANO: For mortality, there is a more explicit tradeoff with loss of reliability and loss of discrimination associated with reporting those outcomes separately. So, we can try to estimate that.

There is already some concern about the signaled noise ratio being a bit on the low side at 30 percent. So, I would expect that that would further decrease when we split the measures.

Again, the other things is, from the consumer's perspective, of course, the consumer is not making the choice about what type of procedure to have. So, I am not sure. That argument is often made, that when there is a choice that is made by the surgeon as far as the particular technique, that that is not appropriate for public reporting because it is not the consumer's choice.

MR. KRESOWIK: This is Tim Kresowik.

I just am going to have to disagree with the idea that the consumers aren't making this choice in reality. As I said before, consumers want an endovascular repair, if it is feasible. So, in a sense, the decision is really, yes, it is a surgeon decision, but to be based on the technical characteristics of the aneurysm and whether it is feasible or not.

I would just like to add one more thing. In this discussion, and I would urge AHRQ to do this, too, I think there is an awful lot of focus on statistically-significant differences in thresholds and not enough on clinically-significant differences.

In other words, from having reviewed a number of these volume outcome studies, when you look at -- you know, a threshold is picked because it meets some statistical difference when, in fact, the difference between two, if you will, volume groups are the difference between 2 percent

mortality and 3 percent mortality or even 2.5 and 3 percent. I think that has to be taken into consideration before too much stock is put on a study that suggests that there is a statistically-significant difference, when, in fact, the clinical differences in terms of mortality are minimal.

So, I would urge the group just to consider all those factor when they look at the data supporting these.

Thank you.

CO-CHAIR TORCHIANA: Thank you.

So, let's I think move on from these two measures and go to 1523, in-hospital mortality following elective open repair.

MEMBER HALPERN: I am going to do these together because the EVAR one and the open one are very similar.

So, as presented by the sponsors previously, the reason they chose to look at small asymptomatic aneurysms is because from many studies the threshold for operating on

aneurysms is 5.5, and that is based on the rupture risk versus risk of surgery.

For women, it is slightly smaller than men. And that is why they chose the thresholds that they did.

When you are operating on a smaller aneurysm, there are reasons to operate, growth rate and, quote, "symptoms", although they are not ruptured. And there are some studies that suggest that smaller aneurysms in a certain subset of patients, such as patients with poorly-controlled hypertension and COPD rupture at smaller rates. So, those are the reasons to operate on smaller aneurysms.

But, if you are going to do that, you have to make sure that your mortality is also low. And as suggested by the sponsors, that means that you have to have done some kind of risk stratification before you even perform the surgery. And, thus, it should be implicit that you have already done that, and

the patients are pre-risk-stratified. And that is why they felt they should just look at those patient populations.

The comments that came up during the discussion of this procedure is the in-hospital mortality versus a 30-day mortality for both measures. Particularly, it came up with the endovascular because a lot of endovascular patients are home within 48 hours.

And as with the discussion earlier with the ambulatory surgery centers, your responsibility doesn't end when the patient leaves the door. And so, there was a lot of discussion about asking the sponsors if they can make it a 30-day mortality instead of just an in-hospital mortality.

CO-CHAIR TORCHIANA: Other comments?

(No response.)

Could the developers respond to that request?

MR. KRESOWIK: Certainly. And again, the intent, I should say the long-range intent, with these measures is to make them universally applicable in terms of being able to capture this data with administrative, or capture the outcomes anyway with administrative data alone.

And that is the real problem with 30 days, is that, obviously, that is the standard. That is ideal. But it requires clear participation in some sort of, if you will, prospective followup with calls, with all those things that, obviously, are the ideal, but are likely never to be universally applicable. So, that is the reason for focusing on the in-hospital mortality.

It certainly would say that in both cases, regardless of the discharge within 48 hours for endovascular, you still are picking up the vast majority of adverse outcomes. And it is just based on numerous studies, that ratio between what happens

within the hospital versus the 30 days is pretty consistent. So, they do track very well together.

And it is certainly our belief that the tradeoff in terms of applicability of the hospital mortality versus the 30-day is well worth it in terms of the burden of data collection and the applicability.

Thank you.

MEMBER HALPERN: Some of the comments were that, since many folks are using NSQIP, that the data could be picked up that way or some other registry, especially with our earlier -- I don't think you were on the phone call earlier, but there is a measure out there to look at the percentage of people participating in some kind of multi-center database.

MR. KRESOWIK: Absolutely. I mean that is where it has got to start. I am actually the clinical lead for NSQIP in our institution. It does not currently capture

aneurysm size.

And I would say, you know, yes, there is NSQIP out there, but it is certainly not widespread in terms of the entire country. And so, again, in order to move eventually to something that could truly be publicly reported at every single institution, that is the reason to kind of design it the way it has been at this point.

Part of the intent is to try to create CPT-II codes that would allow the aneurysm size to be reported administratively, that would be the long-term goal for this set of measures, again, to allow universal applicability.

MEMBER HALPERN: The comments that are coming up here are, then, why not a 30-day mortality code?

MR. KRESOWIK: Again, it depends on what your ultimate goal is. See, the beauty is if you can -- I will try. You are still requiring some sort of voluntary

participation and reporting. Even though that is a laudable goal, we are, clearly, if you look nationally, very far from that.

The opportunity is, again, to be able to use existing administrative data that would not require any participation on the part of the provider, once you have this aneurysm size as a criteria.

So, I will share sort of the possibilities for the future. It is that you could, in fact, if you don't get the kind of reporting that you want, eventually say that the assumption is that, if you are not reporting this as a larger aneurysm, greater than the threshold, the assumption will be that it is within that threshold.

So, again, the idea is to finally get to the point where we can have valid comparison data. We just went through a discussion about all the problems with the AHRQ measures that are existing out there, and this would be something that I think could

actually get us to something that would be scientifically and clinically accepted by the surgical community and get to what we really want, which is real outcome data that patients and others could use.

I would also say that the beauty of these measures, too, is you can eliminate the volume threshold issue. If you are only doing a few aneurysms a year, you had best have a zero percent mortality or else you are going to look pretty bad.

So, with that kind of approach, if you will, you still have a volume incentive, but it becomes one where, if you are really a low-volume provider, you almost have to have a zero percent mortality or you are at least going to be identified, if you will.

So, that is, I think, some of the advantage to this approach. But, again, there is no problem with the 30-day other than it really limits the universal applicability down the road.

MEMBER WILHOIT: A piece of fine print is that the timeframes, again, are written very confusingly. We are talking about that it is in-hospital, but that is not what the document says.

But my real question is the denominator description for these includes people with small aneurysms, but the data that is reported includes people with larger aneurysms. And I am just trying to understand where the data comes from, if the measure only includes people with small aneurysms.

MR. KRESOWIK: Yes, and probably the term small is not, I mean it is a difficult area to try to characterize. The intent with the thresholds here were to, if you will, exclude from the denominator patients with large aneurysms. And that is probably a better way to look at it.

The idea being that, you know, again, where you draw that exact line, the data certainly support that observation is

reasonable, the reasonable alternative for men with 5.5-centimeter or less aneurysms and women with 5.00centimer and less aneurysms. Again, because of less limited data, the validity of that 5 for women is a little bit less robust.

But, anyway, so it is clearly safe to observe those patients. And on the other hand, we clearly know that someone that comes in with a 10-centimeter aneurysm, which we all see, even if they do have significant comorbidity, the high rupture rate of that in the short-term may justify an intervention.

So, the intent really was to exclude from the denominator aneurysms for which this concept of risk adjustment on the part of the surgeon weighing the risks and benefits becomes less important by excluding those. So, that is the intent, not to really identify the use of this term small, if you will.

MEMBER WILHOIT: Okay. No, I

understand that. And that is how your denominator is defined. But in 2D5, in the testing results, there are results provided for men with aneurysms larger than 6 centimeters and women with aneurysms larger than 5.5 centimeters. And I wasn't sure where that data would come from in this measure because those people should have been excluded, according to the denominator definition.

MR. KRESOWIK: And I am sorry, but that was supposed to be the denominator definition. And if there is an error in the submission -- the denominator definition was supposed to be 6 centimeters or less for men and 5.5 or less for women. That would be the denominator inclusion criteria.

MEMBER WILHOIT: And, then, that is what is in the description, but, then, the data that is presented is different than that. And so, I am not sure when the sample sizes are given, do the sample sizes include people

with aneurysms that are larger than those criteria? You know, I am not sure how to interpret the data because the data that are presented, both in terms of sample size and results, include larger aneurysms than are included in the denominator.

MEMBER HALPERN: That data that you are looking at is the volume data. And I think what he was trying to say is that they want to use the volume data. They want to make it an easy reliability.

The data that is evidence for smaller aneurysms having less rupturing is a couple of a very large studies, one out of the UK, that looked at following aneurysms. They basically followed aneurysms and saw what the rupture rates were.

But that is not the data that you are looking at right now. That was just mortality by volume. That is actually the AHRQ data.

MEMBER WILHOIT: But are these the

results -- I guess I am really confused now. I thought the results in 2D5 were the results of the indicator that is being described. Are the results in 2D5 the results of something else? That is what I am trying to understand.

MR. KRESOWIK: I am trying to catch up with you here. I am sorry. So, which measure are we talking about?

MEMBER WILHOIT: I think --

MEMBER HALPERN: It is both of them.

MEMBER WILHOIT: -- it is both of them.

MEMBER HALPERN: Because they have the same background data implied in both. If you look at 2 -- let's see if I can get to it -- 2D5 --

MEMBER SEARS: And you are talking about 1523?

MEMBER HALPERN: 2D5 says refinement of HCUP quality indicators. They are essentially using the AHRQ data, but I

think that is not really the background for this study, this indicator, rather.

MEMBER WILHOIT: So, then, if we don't know the sample size and we don't know the results, are we even in a position to make a decision?

MR. KRESOWIK: I am looking, and I think I am caught up with you now.

So, the testing result was done by the Northern New England group, okay? And so, they are quoting just what the actual mortality was in the thresholds that we described. So, those should be the same thresholds that are in the measure. If there is not, there is an error. So, we are talking about -- so, this is not AHRQ data -- this is actual data from the Northern New England, their registry.

MEMBER WILHOIT: Okay. So, if we are looking at 1523 then, in 2D3, it describes 1201 patients. Is it 1201 patients with aneurysms of all sizes or is it 1201 with

aneurysms below the size we are describing?

And it makes a difference. Because are the data on the smaller aneurysms, is it based on 100 patients or 1,000 patients. We don't know if it includes both.

MR. KRESOWIK: I see what you are saying.

The whole sample is all patients undergoing open elective repair. I am sorry, I don't the breakdown of that test. I was not part of that testing group. And I can't answer the question because I don't have the testing data.

I would assume, okay, just based on other studies, that the vast majority of the patients are going to fall into what we call -- and I don't like this term, either -- the small volume, so the less than 6 centimeters with men and less than 5.5 centimeters with women. Those are the vast majority of patients being done. I assume that is the same for Northern New England, but

I just don't have the exact numbers. I am sure that could be provided for you, though.

MEMBER WILHOIT: And partly, it makes a difference I think even in terms of the relevance of the measure because this is seven years of data for about 10 centers. And if there's only 1200 patients over seven years, and if some of those 1200 patients don't even belong in the nature because they have larger aneurysms, again, that affects the relevance for me.

MR. KRESOWIK: But aren't we confusing two things? The relevance and the scientific evidence for the design threshold is based on large trials. This section we are talking about here is only a testing of the measure more focused on feasibility. I think we are talking about two separate issues.

This has nothing to do with the scientific evidence behind the measure or the choice of volume threshold. This was just, again, to meet the NQF criteria that measures

for endorsement have to be tested in some way, and just to show that you can collect the measure, it is feasible, but that is quite a different question from the scientific validity.

MEMBER HALPERN: Also, you are talking about importance. Okay, the importance is that, as the sponsor just indicated, there's a lot of people getting done at smaller aneurysm size. There is evidence that says you don't have to operate on everybody with a smaller aneurysm. Many of them stay the same and can be watched and may never need an operation.

So, it is important because you want to make sure that people who are getting operations are really the correct people, and that people are making the right choices in operating on those patients. So, if you have a 5 percent mortality in a small aneurysm patient, well, that is actually higher than or equal to their rupture risk.

MEMBER WILHOIT: No, and I understand that, but if a center is doing 10 cases in 10 years, the meaningfulness of the measure I think is limited. And again, I don't know because I don't know what the numbers are here.

MR. KRESOWIK: Can I just take you through that a little bit because I am confused by that?

But if a center is doing, let's say they are doing 10 cases a year. And if three of those patients have mortality, and in this measure you pick that up, that is a 30 percent mortality, which would certainly be something that I would think that people would want to know.

And the point of this is, if you are talking about this population, this should be one that should have an extremely low mortality in order to be, if you will, a justified operation in the first place.

So, I think this is important and

it is very relevant regardless of the numbers that are being performed.

MEMBER WILHOIT: Right, except for when you get very small numbers, there is just such year-to-year variation; it is really hard to draw conclusions or take action.

You know, if you have one complication this year and then none for five years, if the numbers are small, it can just cause such variation from year to year; it is hard to interpret.

MEMBER HALPERN: I think it is similar to what we were talking about the ambulatory surgery centers, though. If you have had a .2 percent mortality and, then, all of a sudden, the next year you have a 30 percent mortality, it may make you go back and look at your process. Or did you get somebody who is -- you can't blame everything on the patient. So, you have to go back and look at your process if you suddenly have a spike in your mortality.

CO-CHAIR TORCHIANA: I guess the point here, it gets back to the study that shows that zero mortality is predictive only of low volume, and that it, in fact, predicts average to high mortality in subsequent years in surgical patients. And so, the whole notion of when is low volume too low is tough.

This seems like a very innovative measure in the sense that it, I think very concisely, pares down this complex issue of risk adjustment using administrative data. So, it is a very intriguing measure. It has potentially got applicability in a lot of other surgical areas, I would guess.

And, you know, these concerns about numbers notwithstanding, I guess the question that seemed obvious to me at first, and now seems less obvious, is, right now, do you actually have -- and I am asking the developer -- do you actually have a way of identifying the small aneurysms in administrative data? And if not, how are you

going to do it?

MR. KRESOWIK: No, and that is the reason why this is being limited at the current time to registry that would include that variable. The Northern New England registry does include aneurysm size as one of their data points.

CO-CHAIR TORCHIANA: But could I ask, then, does the Northern New England registry include 30-day mortality?

MR. KRESOWIK: They do. And again, maybe it is better to just discuss the whole history of this. This measure or this group of measures was originally proposed to go forward as a non-registry measure. You are up against -- and this is a problem with NQF right now -- is you have got this whole concept of testing. Well, you run into the feasibility of how do you test something without creating a way for testing.

So, we have got the CPT-II codes to allow this, are working their way through

the CPT-II, it is called PMAG, so that group. We have also had discussions with CMS about, because you need to partner, if you are going to do this with administrative data, you need somebody that has that administrative data.

So, that is all part of the long-term strategy, but in order to get these considered, so that at least this whole discussion -- and I think you just made a very important point, is this could have applicability in so many other surgical areas. It is sort of what we have all been looking for, is a way to get real outcome data without a tremendous burden of data collection.

So, all those processes are in place. The reason they are here before this panel today is to start getting this concept out there. And it is currently feasible to do within a registry that would collect this data. But this is only the stepping stone, if you will, to hopefully a bigger picture.

CO-CHAIR TORCHIANA: Could I

suggest, though, that when you do this in the registry, assuming we vote it up, that you do the 30-day mortality, in addition to the in-hospital mortality, just to establish the credibility of your assertion that all the deaths occur, by and large, in the hospital?

MR. KRESOWIK: Absolutely. Yes, and that data, that is, the Northern New England does collect the 30-day mortality. And this was just sort of -- that is why I went through the history a little bit.

This was originally designed, and we had hoped to go forward with the bigger picture in mind, but, then, they got pared down, if you will, to just the registry. But Northern New England does collect that data and end point.

CO-CHAIR TORCHIANA: Great.

MEMBER CIMA: I am a little confused now. What are we voting on in the sense of, do you have to be a member of the SVS? I mean this is like the discussion we

had with the STS at the last meeting and this morning.

You know, I was going through here, and it is talking about the New England and the SVS. You have to be a member of a registry, it says. So, is this a pilot or is this for the whole country to do it?

Because the STS covers 95 percent of cardiac surgeons, they said, in the country. What does the SVS cover in the registry?

CO-CHAIR TORCHIANA: I think the registry in this case is in NNE, which is 10 hospitals.

MEMBER HALPERN: There is actually an SVS registry that has started in the last, I think, year and a half that people are starting to participate in. So, there is a growing SVS registry, but it is kind of in its infancy.

However, if the CPT-II codes go through, and notwithstanding the sponsor's

concern about 30 days, I can't believe that it is that hard to collect 30-day mortality because I think most hospitals and places are tracking that now.

But it is the only additional information you need is size of the aneurysm, and it is a threshold. So, once you have above or below that threshold in the CPT codes, and, then, you have your 30-day mortality, it is a yes/no question, alive or dead. So, it should be fairly easy for anybody to do, once those CPT codes are through.

MR. KRESOWIK: Yes, and I think that there is no question that -- I will try to answer the question about whether it is a pilot or not. And again, the intent, the original intent, was to have this go forward more to get this concept endorsed, so that further, if you will, implementation strategy could go on. But we do have the ability to do this right now within the registry.

The SVS registry right now is primary the carotid registry. So, it is applicable to those two measures which are related, but that is not what we are talking about right now.

The Northern New England registry -- and, again, I don't have the -- it is 1213 institutions. They basically cover most of northern New England. So, that is ongoing.

The only reason that you are talking about this as a registry-only measure is because that is the only way we could get this to this level at this point in time. That was an NQF decision. There were multiple phone calls and discussions that suggested that it go this route.

Because, right now, there are a number of groups out there like this one with, I think, a creative approach, but you are blocked by you can't the measure tested until you get it endorsed. You can't endorse it until it is tested. So, it is kind of a

Catch-22 that many groups are in.

So, this is an attempt to try to get the concept which it is feasible to do within the registry setting, but, again, this is just the short-term.

I hope it didn't add more confusion, but --

MEMBER MORTON: I just want to clarify one point. We are talking about a measure that we can only obtain through a registry because you need the anatomic data about the size of aneurysm. And we currently don't have a registry other than the 10-hospital group?

MEMBER HALPERN: You could get that data because every operation somewhere you have a CT scan that has the size of the aneurysm. So, you could get that data, but it would be more laborious to do that if you don't have an easy way to do it, which it sounds like why they were trying to create the CPT-II codes.

So, the question that would come up is, how far along are those CPT-II codes to make this a universally-applicable measure?

MEMBER WILHOIT: And the other thing is the measure is written very specifically in terms of being based on registry data. Switching to CPT-II codes is dandy, but it is not what this measure is.

MEMBER CIMA: Can the staff tell us what is going on here? I mean we have never had one that was just for 10 hospitals. So, why are we voting on this? We are supposed to be here to do an overall thing, right, for the country?

MS. BOSSLEY: Right. So, let's step back for a second and just talk in general because this applies to both the measures that come from STS as well as these.

They are specified using registry data. As they are written, they are not specifically calling out any specific registry and tying it to that. They just happen to be

using data to support for testing and to show use with STS and, also, with New England and the SVS registry.

So, any measure you have before you that uses registry data can be used by anyone who has a registry, assuming they are in that field.

Helen, do you have anything to add?

DR. BURSTIN: Yes. The only other thing I would add is that they could, I mean even the registry measures we have already endorsed have sufficient specifications in them that anyone could take them and apply them doing a chart in your hospital, or whatever the case may be, and, then, look toward the benchmark and see where you are, because we are trying to get to national metrics.

So, I guess the question would be, as this is written, can you take this, with the specifications in this form, tested on a

registry, which is fine -- that is not an issue for us -- and actually use it in a widespread national scale?

MEMBER HALPERN: Again, in my mind, you wouldn't even need a registry because there are only two data points that you need. You need above and below a certain size, and you need alive or dead at whatever time point you say. So, if you say a 30-day time point, I mean you can get that from a chart review. So, in my mind, you don't need a registry to have this be a measure.

MEMBER WILHOIT: You may not need it, but, then, the measure needs to be rewritten, because the measure is defined in terms of a registry.

MEMBER HALPERN: But, then, the same thing applies to all the STS ones we looked at.

DR. BURSTIN: Actually, those are generally written as a real numerator and a denominator, not specifically tied to the

registry. I mean they are elements within the STS registry, but --

MEMBER CIMA: It says "a registry that includes...," and, then, it goes on to specify the Society of Vascular Surgeons quality registry and the vascular surgeons group have such information.

So, either you change the wording --

MEMBER HALPERN: They say any registry, though.

MEMBER CIMA: Yes, but, I mean --

MEMBER HALPERN: It doesn't say the SVS registry.

MEMBER CIMA: -- but it does specify two registries. I am just saying I am not disagreeing with the two points that you are getting at. I am just saying Melinda has said multiple times we vote on what is written in front of us. And what is written in front of us, the way it is worded says using these registries.

MS. MURPHY: Actually, I don't think it says that.

MEMBER HALPERN: I don't think it says that, either. It says --

DR. BURSTIN: He is absolutely right. Under 2A3, it specifically says that the numerator details is a registry that includes hospitalization details. So, that probably needs to be addressed.

MS. MURPHY: But it says what the registry includes. It doesn't say this registry. It has got to be a registry that includes those elements.

MS. BOSSLEY: I think it could be easily reworded --

MEMBER CIMA: Yes, it just a wording thing.

MS. BOSSLEY: -- and taken care of, yes.

CO-CHAIR TORCHIANA: Okay. So, my head is spinning, to put it mildly.

(Laughter.)

MR. KRESOWIK: Mine, too. Mine, too.

CO-CHAIR TORCHIANA: Let me see if I can put together where we stand on these two measures, which are obviously, more or less, one and the same except for the method of repair.

We need to allow activity to occur to move this forward. On the other hand, the way that this is exactly specified doesn't seem to move it forward very much in terms of having a significant impact on a lot of patients at a lot of institutions.

But this is definitely a baby that I think we want to nurture. So, the question is, how do we nurture this measure in a constructive way, staying within our pretty constrained rule set?

MS. MURPHY: And this could be heresy, but one of the ways that it could be nurtured is by continued implementation and bringing it to be considered for endorsement

at a point after some of the questions that have been raised have been answered in the application of the measure.

CO-CHAIR TORCHIANA: So, does that mean we just don't vote at all?

MS. MURPHY: No, I don't think so.

CO-CHAIR TORCHIANA: Okay.

MS. MURPHY: I think the group has to vote in terms of whether or not they believe that at this point in time the measure is matured, tested, and can be appropriately applied on a national basis.

And I am looking in the direction of my partners in crime here.

CO-CHAIR TORCHIANA: And, then, just hold off on the final vote, on the last vote?

MS. MURPHY: Well, in terms of the vote on endorsement, we are going to hold off on it anyway. We will not at this point vote on endorsement.

CO-CHAIR TORCHIANA: Okay.

MS. MURPHY: We vote on the extent to which it meets each of the criteria.

CO-CHAIR TORCHIANA: All right. It sounds good to me.

Does everybody have that? We are going to vote on the criteria, but not on endorsement.

So, 1523, yes/no, important to measure?

(Vote.)

Eighteen, yes; three, no.

Scientific acceptability of measure properties, 1 through 4.

(Vote.)

Two, completely; sixteen, partially; two, minimally; one, not at all.

Does it meet NQF criteria for usability, 1 through 4?

(Vote.)

Four, completely; eleven, partially; four, minimally; two, not at all.

NQF criteria for feasibility.

(Vote.)

Four, completely; ten, partially; three, minimally; four, not at all.

Then, we will abstain on voting on endorsement.

MS. MURPHY: The next vote is whether we agree that it meets all the criteria. That was what the next vote would be, is whether it meets the criteria for endorsement, not --

CO-CHAIR TORCHIANA: Okay. So, we should have that vote then?

MS. MURPHY: Yes.

CO-CHAIR TORCHIANA: Okay. Does it meet all the NQF criteria for endorsement?

(Vote.)

Nine, yes; eleven, no; one, abstain.

Okay. Should we go on to the endovascular?

MR. KRESOWIK: This is Tim.

Can I just ask -- I wasn't certain

what we were expected to get back with. Is that going to be clarified at some point? I thought, prior to the last discussion, there was something the developer needed to provide. Can somebody clarify that at a later point?

MEMBER HALPERN: I think it sounded like, and I can be corrected, but it sounded like one was rewording what your denominator is, to not make it sound like it is so specific to the surgery, those vascular surgery registries.

Two is the 30-day mortality issue.

And I don't know. What was the other thing?

MEMBER WILHOIT: Providing data that gives results for the measure, including the denominator and what the numbers are. You know, what the sample size is or the population for the facilities being measured, but real data that exactly measures what is in the measure.

MS. MURPHY: And for all of the

measures, we will send information back to the developers about the additional information that the group is requesting. So, we will get it back to you, also, written.

MR. KRESOWIK: Thank you.

CO-CHAIR TORCHIANA: I think some notion as to what the time course might be on the CPT-II codes as well was raised.

So, let's follow through now and vote on EVAR. It will be interesting to see how these correlate, if anyone can remember how they voted on the first one.

Importance to measure.

(Vote.)

Twenty-one, yes.

Scientific acceptability.

(Vote.)

Five, completely; thirteen, partially; three, minimally.

Usability.

(Vote.)

Three, completely; fifteen,

partially; two, minimally; one, not at all.

Feasibility.

(Vote.)

We can't seem to get that 21st one in.

Five, completely; ten, partially; five, minimally; one, not at all.

Does the measure meet all the NQF criteria for endorsement?

(Vote.)

Nine, yes; twelve, no.

So, I would suggest we are at the time for our break. So, should we break now and pick up these right where we are?

MEMBER SEARS: This is Nick Sears. I have a time constraint. I have to leave for the airport in about 35 minutes.

MR. KRESOWIK: This is Tim, too.

We just have the carotid ones. Do you think we could deal with those? Because the discussion should not be substantially different than the one we just had.

CO-CHAIR TORCHIANA: Okay. I thought you were going to say should be short.

(Laughter.)

MR. KRESOWIK: Well, I meant that. I meant that. There shouldn't be any new discussion, but I --

CO-CHAIR TORCHIANA: Okay. So, well, let's see. So, we have these three remaining vascular surgery measures. Can people take trying to knock those off before the break?

Okay. So, let's do 1548 -- this is the surveillance one -- first.

And who is commenting on this?

MEMBER KENNEDY: This is Eileen Kennedy.

CO-CHAIR TORCHIANA: Hi.

MEMBER KENNEDY: I am going to do this one.

Hi.

Okay. So, yes, Measure 1548, surveillance after EVAR. And as stated

previously by the sponsor, this measure calls for one followup imaging study after three months and within 15 months of the repair.

It is important to note that there is a risk of potential failure of the endograft therapy with problems leading to potential rupture and increased mortality. These complications may be identified during optimal surveillance with the use of imaging scans.

And there were opportunities for improvement identified within the measure material.

The measure does not require risk adjustment.

The Work Group was divided on the degree to which this measure meets the scientific acceptability criteria. It was noted that there are currently the two registries that are used to record the surgical details the EVR procedures. However, the team raised a similar question to that of

the previous measures, to the requirement of membership in these registries in order to collect the surveillance data.

There was also an additional observation made on the potential socioeconomic impact of patients not complying with the required followup due to the cost of the scans or necessary travel.

And, then, the final issue that was raised was just around the use of this measure for public reporting.

CO-CHAIR TORCHIANA: Other discussion?

MEMBER HALPERN: Yes. I think the issue of public reporting was one of the major issues we discussed. Because what does this mean? There are so many reasons why a patient may not have a CT scan.

The idea thing is that you have to have some kind of followup for your endovascular repairs because there are various types of endo leaks that occur, and some of

them may lead to rupture.

As stated by the sponsor, the exact timing is kind of still being a little bit debated in the literature and, also, what kind of imaging modality you need to do as ultrasound becomes better. The sort of original standard was x-rays and CTs, which is an awful lot of radiation over time.

And so, it wasn't quite clear how this would be used and in what manner it would be used, and what its importance would be.

MR. KRESOWIK: So, let me just clarify a couple of things.

The numerator for this is not CT scans. It is any imaging modality, which would include duplex.

And I can say that, as I tried to say before, even though there is some controversy or question about what is the ideal post-procedure imaging strategy, anything that is out there right now would suggest that this is a floor. Okay? So,

anything less than this would not be possibly acceptable.

And, you know, the problem is you have got to start somewhere. I think you would get into more controversy if you tried to have a more aggressive imaging strategy than this. But I don't think there is anybody or any evidence that would support anything less than this measure.

So, this should be looked on as an absolute floor, and it does not specify any particular imaging modality. So, duplex, ultrasound, which avoids the radiation risk, would meet this measure.

MEMBER HALPERN: And what is the public reporting of this going to be? How do you see that being utilized?

MR. KRESOWIK: I guess I don't see it any different than any other process measure. It is a process measure, just like lots of process measures that are out there. It has the same purpose.

I mean the ultimate purpose is to get people to do the right thing. But I guess I would view it as any other process measure.

CO-CHAIR TORCHIANA: And how is this data acquired outside of a registry?

MR. KRESOWIK: Well, I mean, again, I don't want to reopen the whole box we just went through. But, again, it was because of the limitations of getting it to this point, it had to be constrained within a registry format. It does not, as was said before, does not require the specific registries that were cited to support the measure. But, clearly, this data could be done simply with administrative claims, but we couldn't test it except with what we had in the registry format.

So, I hope I haven't added to the confusion again, but this is just the advice and the limitations of the current process getting it to this point.

CO-CHAIR TORCHIANA: I think we

get that.

Any other discussion?

MEMBER SAIGAL: Most of these ruptures happen after six months? They don't happen early, is that right?

MEMBER HALPERN: Yes, for the most part, and rupture is actually relatively rare, but endo leaks are not.

MEMBER SAIGAL: So, the evidence on imaging happens after six months?

MEMBER HALPERN: Different people have different criteria. Some people image at three months. Some people image at six months. It also may depend on how comfortable you were with your repair at the time of the procedure. Like was it an angulated neck where the risk of endo leak is higher? You know, did you see an endo leak, type II endo leak, at the end of the case?

So, there might be other things that would promote you to have an earlier, rather than a later, imaging. But he is

correct that anything longer than this would really be unacceptable in any endovascular repair.

MEMBER SIPERSTEIN: Can you clarify for me the process of a patient has their procedure at one institution and has their long-term followup in a remote state or country, how that is dealt with?

MR. KRESOWIK: Well, again, the idea is that whoever -- if you are looking at this in a registry, you would be required to know that. You are still the responsible person.

How it would act in reality, this is not uncommon. Let's say a patient had it done in another state, but it would still be -- and this happens even within a state, where you might have the study done elsewhere, but you would be, the surgeon would be responsible to find out what the results of that information were. So, you would have that available to you, regardless of where the

followup occurred. So, this would be the onus on the surgeon, if you will, to know that that study has been done and had the information as to the results of that study.

CO-CHAIR TORCHIANA: Okay. Let's vote on this and, then, try to get the last two in on carotid.

CO-CHAIR MORRIS: Before we vote, actually, could I add one more thing? So, our issue with the registry inclusion exists with this measure as well in the numerator. I just want to make sure that everybody is aware of that.

CO-CHAIR TORCHIANA: Agreed.

Importance of measure, yes or no?

(Vote.)

Twenty-one, yes; one, no.

Scientific acceptability, a scale of 1 to 4.

(Vote.)

Three, completely; fifteen, partially; three, minimally.

Usability.

(Vote.)

Three, completely; fifteen, partially; three, minimally.

Feasibility.

(Vote.)

One is not in. Keep pushing.

Three, completely; eleven, partially; five, minimally; two, not at all.

Does the measure meet all the NQF criteria?

(Vote.)

Keep pushing.

Five, yes; fifteen, no; one, abstain.

Thanks.

So, I think the message on that measure is very similar, as was pointed out, to the prior two, in that this confusion around registry versus something that is aimed for a broad use is, I think, one issue.

And, then, the second issue is,

how, short of a registry, will the followup information ever be obtainable? It seems, by definition almost, that it will only be via registry.

So, let's handle 1540 and 1543 together.

Nick?

MEMBER SEARS: Yes, that sounds like a plan.

I think we are going to have the same discussion. 1540 is postoperative stroke or death in asymptomatic patients undergoing a carotid endarterectomy, and 15432 is postoperative stroke or death in asymptomatic patients undergoing carotid artery stenting.

Our group felt, everyone felt that it was a reasonable measure from a measuring standpoint. Some of the issues coming up, as I think we heard from the Society of Vascular Surgery, was the definition of what asymptomatic was. I think a couple of questions were raised around, how do you

document somebody is asymptomatic? Is it CPT references or how can that be documented? And I think that is a fair question, given the fact that patients don't necessarily see the same doctor over and over again.

But both measures, other than that, met scientific reliability.

For the No. 3, efficiency, we didn't have a whole lot of discussion there. Again, public reporting and the registry issue came up.

So, overall, I think most people felt that these were reasonable measures.

MEMBER HALPERN: I would say that the asymptomatic is actually easy with ICD-9 codes because you have to report them together with your surgical code, as to why you are doing the procedure. And it is either a symptomatic or an asymptomatic carotid.

MEMBER SEARS: Okay. Well, that should take care of that problem, then.

MEMBER WILHOIT: Although I think

you would want to test that to see. You know, ICD-9 codes aren't always, don't always correlate with the medical record. So, I think don't know without looking.

MEMBER SEARS: Yes, and I would love to just say, yes, that fixes the problem. Unfortunately, the current administrative data doesn't really capture this.

So, it is like the discussion before. The intent is to actually get this definition in the form of CPT-II codes. That process is underway.

But, for the purpose of what you are dealing with today, that definition is part of the existing registries, not only the Northern New England, but also the SVS carotid registry. There's multiple other registries dealing with carotid disease out there.

And I wish I could say all the definitions are identical; they are not. But, unless you go forward with this kind of process of trying to get there -- we do need

to get a common definition. But it is possible to do, and we still need to ultimately get to standardization or harmonization of the definition. And that was the intent, ultimately, the CPT Category II codes.

CO-CHAIR TORCHIANA: Okay. Other comments?

MEMBER WILHOIT: Even though CPT-II codes are a good goal for the long-term, we see in our claims almost no CPT-II codes being used currently. So, it is a good goal. It is a direction to work. But, until they are used consistently, they are not necessarily useful for measurement.

MEMBER HALPERN: I would say, also, that by using codes, I mean you can say that for any of the measures we are looking at. Codes can always be inaccurate.

But when you are charging for your surgery, you have to put a diagnosis code. So, I guess what I am saying is you will have

to trust that the surgeons are accurate or whoever is doing the procedure is accurate in terms of, are they doing it for a symptomatic or an asymptomatic patient?

CO-CHAIR MORRIS: It seems to me that this is part of where testing the measure comes in and looking at the adequacy of testing, and whether or not you are measuring what you are think you are measuring, which if this vote goes like the previous votes, and SVS comes back to us, maybe that is something that they could also include in what they bring back.

MEMBER HALPERN: How accurate the codes are, you mean?

CO-CHAIR MORRIS: Right, if they have registry data available that can compare.

CO-CHAIR TORCHIANA: Any other comments?

MEMBER HALPERN: Like with the smaller aneurysms, I think these measures are important in that, if you are to get benefit

from an asymptomatic carotid, you have to live for five years. And your risk of your procedure has to be about 1 to 3 percent or less, probably closer to 1.

So, if you are doing the procedure, you had better have good results for asymptomatic patients because it is essentially a prophylactic surgery.

CO-CHAIR TORCHIANA: Could I ask the developer, what percentage of the carotid stents are done by vascular surgeons versus interventional radiologists versus cardiologists? And what the sort of registry and coding practices of those other specialties? Are they similar or are they very disparate?

MR. KRESOWIK: Yes, I can't answer that question. It is a continued evolution. Just it is an unanswerable question, I think, at this point.

MR. ANDERSON: Excuse me. There are some data on that.

MR. KRESOWIK: It depends on the hospital, you know.

MR. ANDERSON: Yes, I am Skip Anderson. I am a cardiologist with the ACC.

And we reviewed that material last fall, and for the performance of procedures, carotid standing procedures, it is about a third done by vascular surgeons, about a third by radiologists, and about a third by cardiologists.

CO-CHAIR TORCHIANA: I guess I would suggest to our colleague from the Society of Vascular Surgery that it is really important that the measure for carotid artery stenting be developed in conjunction with those other specialties. I would think that is pretty critical from the standpoint of the NQF endorsing them.

MR. KRESOWIK: Yes, and we have had those discussions previously. I mean those discussions about this measure have occurred.

MEMBER STAFFORD: It is Renae Stafford.

I would agree with that. Because looking at the two measures, the level of analysis is different. So, for the carotid endarterectomy, the level of analysis is at the clinician and group level. And for the stenting, it is at the facility and agency level. And that doesn't make any sense to me.

MEMBER SEARS: I thought it was at both levels. I thought it was lifetime for the surgeon and annual for the hospital, reporting both.

MEMBER STAFFORD: It is hard to tell. Looking at least at what the summary that we received from the NQF, the level of analysis is different for the two of them. I haven't gone to the specific document.

MR. KRESOWIK: They shouldn't be, but if that is, it is an error.

MEMBER SEARS: The only other question I had, because I don't do stenting,

but is there any risk for the people who do stent in following patients longer than 30 days? Because can the stent dislodge or do what the abdominal stents do? The 30-day numerator number may be off for the stent. I don't know the answer to that.

MR. KRESOWIK: In both cases, I mean, yes, there can be late complications. They are not substantially different. In the studies that have been done in comparison, the kind of long-term outcomes are pretty similar. Obviously, there are differences in the procedural time, but, yes, there can be late -- it is not stent dislodgment in this case. You know, you can always have a late embolism or thrombosis, but they are not really actually different between the two procedures.

MEMBER SEARS: Okay.

MR. KRESOWIK: And they don't require -- I mean the surveillance approach would be similar.

CO-CHAIR TORCHIANA: Okay. Let's, then, vote on 1540, a stroke or death after carotid endarterectomy.

Does it mean NQF criteria for importance?

(Vote.)

Twenty, yes; one, no.

Scientific acceptability.

(Vote.)

Six, completely; fourteen, partially; one, minimally.

Usability.

(Vote.)

Five, completely; fourteen, partially; one, minimally; one, not at all.

Feasibility.

(Vote.)

Four, completely; thirteen, partially; three, minimally; one, not at all.

Does the measure meet all the NQF criteria for endorsement?

(Vote.)

Thirteen, yes; eight, no.

Now the same questions on carotid stenting.

Importance to measure and report.

(Vote.)

Keep pushing those buttons, please.

Twenty-one, yes.

Scientific acceptability.

(Vote.)

Six, completely; fourteen, partially; one, minimally.

Usability.

(Vote.)

Six, completely; thirteen, partially; one, minimally; one, not at all.

Feasibility.

(Vote.)

Six, completely; eleven, partially; three, minimally; one, not at all.

Does the measure meet all the NQF criteria?

(Vote.)

Can't seem to get that last one over the finish line. It may be pushing the buttons all these times is wearing the batteries out.

Fifteen, yes; six, no.

Okay. Even though the last one is on carotid, I think we should take our 15-minute break and reconvene at, let's try to shoot for 4:30.

MR. KRESOWIK: Could I just say thank you to all of you for putting up with having this as a phone-in. We really appreciate the opportunity, and it is always difficult, having done this many times.

My only other request, I would be curious if someone there could try to ascertain why the difference in meeting the criteria between the carotid and the aortic aneurysm measures. I think that would be helpful to us to understand some of the feeling in the future.

So, anyway, thank you very much.

CO-CHAIR TORCHIANA: Fatigue, yes.

(Whereupon, the foregoing matter went off the record at 4:19 p.m. and resumed at 4:39 p.m.)

CO-CHAIR TORCHIANA: We should reconvene. I am sorry. We have gone a little over our time.

So, our next measure is 1531, followup assessment of stroke or death after carotid revascularization.

MEMBER ROGERS: `Tis I. `Tis I.

(Laughter.)

Okay. So, this is an interesting initiative that is actually a process measure. It is not an outcomes measure. Really, it is a measurement of how frequently an examination was done within 21 to 60 days following a procedure.

The interesting part is so there is a challenge there that we have talked around and about earlier today on other

procedures where a measurement is requested following discharge.

The added feature to this one is that the exam, then, is requested or required to be done by someone who is certified, NIH Stroke Scale certified. And it is also specified that it not be done by the operator, so that the surgeon involved is not the person.

So, when we discussed this on the phone, a couple of issues came up. One, of course, was the feasibility and the likelihood of being able to retrieve this information subsequent to a procedure being done, particularly because the hospitalization is so short that it just isn't going to happen, then, obviously, in 21 or 60 days. So, it is the issue there, and, also, the issue of certification.

And if there are changes noted in the exam, if it was done by someone certified at 21 days, and the initial baseline exam was

done by someone not certified, then, does it count or does it not count?

So, I think the consensus is that it makes sense. I think it really is a feasibility issue more than anything else.

MEMBER DUTTON: To me, this measure seems like an effort to gather research data rather than quality data. We have had a couple today that have kind of had that flavor, like the reason to have this measure is to generate data for somebody's paper about strokes after carotid.

Is there any evidence here that this will improve patient outcomes?

CO-CHAIR TORCHIANA: Do we have a comment from the developer?

MR. ANDERSON: Yes. I am sorry. Could you repeat the question? What was the --

MEMBER DUTTON: This is a process measure, as mentioned. I think to have a process measure for quality purposes, it has

to be strongly linked to an outcome. So, how does doing this additional assessment improve patient outcomes?

MR. ANDERSON: Well, it is part of obtaining outcomes, outcomes assessment. It is an independent neurological exam based on the NIH SS, Stroke Scale, which is a relatively-simple clinical tool which many people can use to try to find out if a neurological event has occurred.

And the idea is to be able, for internal quality purposes at institutions or for reporting, public reporting, to try to track clinical outcomes, particularly neurological outcomes. But if you don't have a uniform, standardized tool for assessing that, then it becomes a little bit problematic, and you have institutions that don't do it at all or that aren't obtaining quality data.

So, the idea around the tool, the simple clinical tool, is to provide a little

bit of a standardization of assessments.

MEMBER HALPERN: Another question that came up along those lines -- and Dr. Rogers sort of referred to it before -- was that, why not do this during the hospitalization since the majority of incidents happen soon after the procedure?

And from a feasibility standpoint, that would also be easier to make sure you have somebody who is not the provider who performed the procedure doing it, because you could have, say, a nurse on the floor. The tool is very simple, because I actually looked at it online after our phone discussion.

MR. ANDERSON: Well, the background behind that was that, of course, patients should have exams before and after the procedure, and the expectation is that they will.

Some clinical events happen after discharge. Although it is oftentimes not reported as such when the trials are

published, most of the clinical trials used a 30-day endpoint. However, a few of them delved into the occurrence of events in-hospital and those after discharge and up to 30 days.

And there were very little bit, but about 10 to 25 percent of the neurological events occurred after discharge, between discharge and 30 days. So, you do miss a substantial number of them if you only obtain hospitalization data.

And since most of the evidence base is centered around 30 days, the proposal was that 30 days with a window, 30 days become an acceptable reporting point.

MS. FITZGERALD: And this is Susan Fitzgerald from the ACC.

I just want to say the specifications are that the examiner should be independent; it doesn't have to be, but we recommend that they should be. It is apparently a more valid test if somebody other

than the operator performs it.

And I think about 90, between 96 and 98 percent of the NIH Stroke Scales that come through our registry, the examiner is certified. They are a certified examiner and have gotten their NIH certification.

MEMBER ROGERS: I guess I should know this. Tell me again or remind me where this data is collected.

MR. ANDERSON: Well, it begins at the hospital. It begins at the institution.

MEMBER ROGERS: I know, but where does it end up?

MR. ANDERSON: Yes. Well, for the ones that participate in our registry --

MEMBER ROGERS: Right.

MR. ANDERSON: -- they accumulate the data and transmit it to the ACC.

MEMBER ROGERS: So, it raises the issue of belonging to a registry once again. And can you tell us what options there might be in addition?

MR. ANDERSON: Well, of course, we would like for all institutions to collect their own data and engage in quality improvement initiatives and do public reporting on their own without participating in registries. But, you know, the trend has been to do at least regional-based or sometimes nationally-based registries for doing comparisons between institutions, to try to look at variability and quality improvement initiatives on a larger scale. And we support that, too.

But the level of collection is at the institution. We currently have about 178 sites that participate and have about 15,000 records on carotid revascularization. And the data is transmitted to a central warehouse, where it undergoes analysis, and there are reports generated back to the participating on a quarterly and, then, an annual basis for them to review their data.

Of course, we would hope that we

would review their own data anyway, since they are the ones that sent it to us, for internal purposes. But we were able to provide some comparative analysis of the various institutions.

MEMBER WILHOIT: One of the assumptions that was stated in the measure at the beginning is that everyone, obviously, has an office visit following surgery. I think that that isn't always an accurate assumption, No. 1.

And, No. 2, since this requires a visit to someone other than the surgeon, I think that is a really big assumption, that everyone already will have a visit.

Also, the numerator time window, as we have seen on a lot of other measures, doesn't match what is described in the numerator.

It also seemed like there should be exclusions for people who died because people who died wouldn't be available for

followup.

And, also, people with a prior neurologic event or a prior stroke, it seems like there should be some way addressing those people perhaps in the measure as well.

MEMBER HALPERN: Although I think that what they are saying -- correct me if I am wrong -- but my sense of the measure is just that the exam was done, not that there was a change, but just that the exam was done. So, I understand why they don't have necessarily prior exams because they are not making a comparison. They just want to make sure somebody is actually examining the patient sometimes.

But I do wonder, I still wonder about making it 21 to 30 days instead of some time period after the surgery because, you know, most of the practices I have seen, they see their post-op patients a week or two later. It might not be at 21 days and they would fall outside that window.

MEMBER DUTTON: Wouldn't there tend to be a bias in the data? That is, if you had a stroke, you are more likely to go back and see your doctor?

MEMBER HALPERN: Well, you might go back to see a doctor, but it might not be the person who is involved in the procedure or necessarily related to the institution where they came from.

CO-CHAIR TORCHIANA: Could I ask what is a fairly obvious question here? This is the American College of Cardiology measure. And so, I assume we are talking about carotid stenting predominantly and cardiology practice predominantly?

MR. ANDERSON: Yes. No, the registry that we operate is actually a partnership between eight societies, including the American College of Cardiology, the American Academy of Neurological Surgeons, the College of Neurological Surgeons.

So, it is a partnership. All

societies have input, and they all help develop the various measures and reporting standards. So, it is joint effort that is involved in this.

CO-CHAIR TORCHIANA: Is it endarterectomies as well as stenting?

MR. ANDERSON: Yes. We try to capture and encourage reporting on all types of revascularization, both endarterectomies as well as stenting.

And the data that we have accumulated so far is roughly equal between the two types of procedures. Last year, for 2010, there were 2500 endarterectomies and 2500 carotid stents that were submitted to the registry.

MEMBER HALPERN: Is SVS involved in your registry?

MR. ANDERSON: SVS has been involved in several levels. In fact, there are discussions going on with them right now about partnering up, as was discussed earlier,

for this.

In terms of organizational, no, they are not part of, they are not one of the other eight groups that are involved right now. But that continues to be discussed.

CO-CHAIR MORRIS: Can you speak to the 21-to-60-day window that was raised?

MS. FITZGERALD: Well, what we find is that people submit records on followup, and they submit the data followup.

And we understand that some patients have followup at seven days, but if you have a patient who has a followup at seven days and other people are submitting followup at much greater than 30 days, we find we can't compare the two groups. So, we created that followup window to be consistent.

Let me look at a couple of your questions. That was one question.

The other one was that, if you died, you would be counted because you are assessing. You would not expect to have a

stroke scale if you died. If you died in the hospital or you knew that the patient died within 30 days, submitted a followup record, you did the assessment accurately.

And, then, the third thing is prior strokes are not excluded. Evolving strokes and dissection are excluded because carotid endarterectomies and stenting are typically done on an elective basis to prevent strokes. So, if you have a patient who is very sick who is coming in, who they are really not supposed to have procedures like that, but do, we excluded them in the followup because their stroke rates are higher, and their assessment and followup is a little different.

MR. ANDERSON: One other comment is that, for the last decade, the clinical trials that have been published on carotid revascularization have pretty much all centered around a 30-day assessment. And the window is varied between 15 to 40, 25 to 60,

but those numbers were chosen as a sort of average value from the clinical evidence base from the trials that have been published over the last 10 years. The 21 to 60 works out to be roughly a window, an average window, around 30.

MEMBER ROGERS: Can you give us some idea about the relationship of the numbers of reported cases you had last year versus the number, an estimate of the number of cases that are done in the country for the year? So, we are talking 10 percent or 25 percent? Do you have any idea?

MR. ANDERSON: Yes, I don't think we have those. For 2010, I don't think we have an estimate on that.

MEMBER ROGERS: So, the question has been asked by Vivienne, and it has to do with, do you see yourself as a competing entity with SVS, or others, for that matter? And I wonder whether there is a possibility of harmonizing with some of the stuff they do,

perish the thought.

MEMBER HALPERN: I mean that would be the idea, is to have, you know, a registry where we all agree on what the data points are going to be.

MS. FITZGERALD: Right, and there is a physician that is I think the President of SVS that is very close to a physician at the College, and they both work at the same center, who have kind of forged some recent conversations about partnering and moving forward with carotid as well as peripheral vascular initiatives.

So, we did not, when we first launched the registry, we actually had conversations about it, but our registry launched, and so did theirs, and it was kind of too late. So, we reintroduced those conversations.

MEMBER CIMA: I had a question about feasibility. I mean, if you say that it doesn't have to be the same person that does

the procedure, what if they have a procedure and they go see their general practitioner? First of all, you have got to get them to be certified on this scale, the NIH scale. So, you have already made it very difficult to some extent, even though it is an easy thing to do and they can get it offline.

But, then, how are we going to capture the data that they had a carotid here at this major center that they got referred, and, then, the follow up with their neurologist in Boise, Montana, or Boise, Idaho? Sorry.

(Laughter.)

What is the feasibility of doing it? Are you saying they have to be in a registry then? Because you don't want to go down that road because we have been down that road a couple of times.

(Laughter.)

But what are we saying here? How are you going to do the followup? I am just

looking at the usability and feasibility aspects.

MR. ANDERSON: Well, I think, as was mentioned earlier, the responsibility doesn't stop when the patient leaves the hospital. People that are engaged in carotid revascularization have a responsibility to have some sort of followup over what happened to the patient. And that includes a neurological examination.

It is easy enough to transmit reports. And we hope that the responsible operator would have enough of a relationship with referring physicians, should they be located far away, to get some report back, and that at least at the institutional level, that ought to be a relatively-common, quality-type initiative that institutions, hospitals, should be engaged in.

And there probably are cases, some cases, where there may be a distance between the patient, where they get their followup,

and the operator center, but, hopefully, that will be harmonized, especially in this increasing age of electronic transmission of medical record data. So that, getting back a report ought not to be all that difficult. They ought to become expected.

CO-CHAIR MORRIS: I agree with you that people who do these procedures should follow up their patients afterward. I don't think you have really answered the question of, what do the vascular surgeons do who follow up their patients in, say, seven to fourteen days, so they are not inside of this window, but are perfectly responsible and perfectly interested in providing high-quality care.

And, then, I have a second point which I would like you to also address. That is in the 2B and 2C area of this, reliability and validity testing. I am wondering if Group B discussed this in your conference call previously. And if you could clarify a little

bit about those blank places in the reliability and validity testing areas?

So, two questions. One is, can you talk a little bit more about the window? Because if a vascular surgeon follows up between seven and fourteen days, they are not within this window.

And, then, the second part is the testing.

MR. ANDERSON: Yes. Well, again, the window was chosen as a kind of averaging phenomenon. The discussion around that was that seven days was too soon, that a later time point because events do happen was a more preferable time point, and that that ought to become more ingrained and embedded.

You know, we can't force people to do these exams or do this reporting, but the desire was that it should be between 21 and 60 days, based upon the evidence that is out there in terms of clinical trial data for both forms of revascularization and the occurrence

of events after discharge up to 30 days, that the window of, roughly, 21 to 60 days was acceptable.

There are trial data out there that as short as 15 days. I don't know of any that are as short as seven, but certainly 15 days and extending out to around 60, 65 days. But seven was considered to be a little too soon.

MEMBER DUTTON: Well, again, why isn't this an outcome measure? I mean, if you are putting this much precision into measuring every patient in a very precise way, why not report the outcome?

MR. ANDERSON: Perfect. That is an excellent question. And obviously, everybody would like to move that way, toward an outcome assessment, and reporting of outcomes.

In order to be fair, it would have to be a risk-adjusted outcome or risk-standardized outcomes. And we are moving in

that direction.

Risk models do not exist yet. But in order to get to those outcomes and risk-adjusted outcomes, you have to have a process for ascertaining the outcome, and the process is doing a standardized neurological examination. Without the neurological examination being done, you really don't have a fair assessment of the outcome.

CO-CHAIR TORCHIANA: Just a suggestion on the best-of-breed theory, the idea of maybe doing this kind of followup only on asymptomatic patients, where a risk adjustment, as we have learned today, is of minor significance, might be a way to thread that needle.

MEMBER WILHOIT: Yes, and I think that the followup timeframe of the zero to 14 days, it is interesting --

MS. FITZGERALD: I think, as ACC, we think more of the stenting and not as much the endarterectomy in the followup. The

operators performing the stentings, you know, we have talked to them about this, and that seemed to be an appropriate timeframe.

We could drop the front end of it, but we didn't want to compare patients that went into a doctor at seven days with a patient who was going in 45 days with different outcomes.

MEMBER HALPERN: Yes, we tend to see the patients early because we want to see what their neck incision is --

MS. FITZGERALD: Yes.

MEMBER HALPERN: -- especially if they had any liminal swelling or, you know.

MS. FITZGERALD: Right, and that is a really good point. We could drop that end of it.

And, then, the other thing I wanted to mention is, in working with interventionalists and some cardiac surgeons, we always have struggled to report myocardial infarction and some other outcomes after a PCI

or a surgery. And it is because some hospitals report, assess it differently than others.

And we work with a lot of payer programs, and we identify which of our metrics and measures are appropriate to use in a pay-for-performance program or not. And MI is never it because some hospitals have an MI rate of 7 percent because they check biomarkers on every single patients, and others don't. So, we have always stressed the importance of a proper assessment in a patient post-op.

We actually did a really interesting paper on PCIs. Hospitals that checked biomarkers routinely had a much higher-quality hospital. They had a lower mortality rate. They had much higher adherence to process outcomes, medications prescribed at discharge, and stuff like that. It was interesting.

CO-CHAIR TORCHIANA: There is no

doubt that you find problems proportionate to how carefully you look for them.

And could I ask, I think the answer may be fairly straightforward, but the reliability and validity testing, this is obviously, a binary question. Either the exam was done or not. But do you have any testing of this that you could fill us in on?

MS. McGUINN: What we have is performance rates in our registry. So, we haven't audited the data to compare it to medical record, or something like that, to get reliability of the data.

CO-CHAIR TORCHIANA: I'm sorry, I couldn't hear you.

MS. McGUINN: We haven't established reliability by comparing to a medical record or something like that, but we have distribution of performance within the registry.

CO-CHAIR TORCHIANA: Okay.

MEMBER WILHOIT: I just keep

thinking, if I were the patient who had had this procedure done, and it is three weeks later, I have been and I have had my followup visit with Vivienne, and she checked my neck and I was fine, I am a alert, I am oriented, I am back at work, I am functioning normally, I know I am alive, I know I didn't have a stroke, why in the world would I, the patient, want to take time off work to go see some doctor I had never seen to have an exam to see whether I had a stroke? And, then, I have to pay out of pocket because I have a co-payment or co-insurance, or whatever.

I mean that is where I am stuck. It is, why would I ever bother to do that?

MEMBER DUTTON: Yes, is there any evidence that you find more events this way than what the previous two measures find?

MR. ANDERSON: Are you asking about between hospital discharge and 30 days?

MEMBER DUTTON: Sure.

MR. ANDERSON: Yes. Again, based

upon the report the reporting from trial data, it is roughly 10 to 25 percent of the neurological events occur after discharge, up to 30 days.

MEMBER HALPERN: But how long after discharge? So, like, would you catch most of them at seven days or fourteen days, or something that is more reasonable for us who would want to follow up on other aspects of their post-op care?

MEMBER DUTTON: Or would you catch them because the patient says, "I'm not right. I need to go see my doctor."?

MR. ANDERSON: Yes, well, in the trials everybody had the neurological exam. So, for the trial data there's better homogeneity of assessment. Sure, after seven days, it is a declining function. Where you draw that boundary is based on a lot of things, including the ease, and seven days would be okay. I mean the threshold could be set at seven days. That would not be

difficult to do.

MEMBER CIMA: Just to follow up on Carol's point, though, I am looking at it, the value you add to the patient. So, if we have set it at seven days, twenty days or whatever, but how many patients sit at home with facial droop and don't go see anyone? I mean, why are we going to say every single person that has surgery has been seen, has their followup visit, they're global, it is all paid for? Now they have to go make another doctor's appointment. What is the value-add to that? For the patient and for the payer and for Medicare, what is the value-add?

MR. ANDERSON: Well, actually, the discussions that we participated in were actually the other direction. It was, if you were comparing institutions A to B to C to D, and you were looking at their reported complication rates or reported event rates, and you knew that Hospital A was collecting excellent-quality data and had lots of exams

being done, neurological exams, and Hospital B wasn't, how reliable would you feel or how comfortable would you feel as a patient, say, if you or your relative had to go to A or B or C.

So, actually, when the discussions were occurring, it was actually the reverse side of that, looking at it, and how can you compare institutions, knowing that one might be doing a really good job of collecting data and another one might not?

MEMBER HALPERN: I think what he may be asking is, the later exam, how much value? Going to a different provider to get an exam, because your suggestion is that it is somebody other than the person providing the procedure, how much does that improve quality?

MR. ANDERSON: Probably a lot. Most of the neurologists I work with, the examiner is a nurse in the office. So, the followup exam generally occurs, and one of the nurses who are stroke-certified does the exam,

fills out the form, and actually faxes it over.

CO-CHAIR MORRIS: I would like to say something about that. You say that it probably improves the quality a lot, but we don't really know because all you have is a distribution. You haven't really tested it. So, we don't actually know that.

I think that this is very interesting, and your comments about, essentially, the culture of safety and high quality that is reflected by performing this exam, that is a very compelling argument. But it just seems to me that there are parts of this measure that are inadequately developed.

CO-CHAIR TORCHIANA: Okay. Let's go on to vote on 1531, followup assessment of stroke or death after revascularization.

Importance to measure and report.

MS. MURPHY: Before the vote, can I just point out that a part of the scientific acceptability that you have to look at is

testing, and there is no, I mean, there is no information? And we do have the information that reliability has not been tested.

CO-CHAIR TORCHIANA: So, importance to measure.

(Vote.)

One, yes; thirteen, no. Oops, excuse me. Seven, yes; thirteen, no. I am reading the wrong numbers. Okay.

CO-CHAIR MORRIS: As a group, do we want to request that the ACC consider revamping this and bringing it back?

(Chorus of yeses.)

MEMBER ROGERS: If I may make one suggestion, we play in this vascular field in the State of Washington. One of the things that has become very obvious is the tension, not necessarily the tension, but the relationship between and amongst the three specialities who do this work.

I think this is an important measurement, but I think it would be a problem

if we were to bless one professional entity with a measurement that affects three different specialities.

So, my invitation -- it is not coming from the rest of the group -- is that you have more, hopefully, productive discussions with the other disciplines. Because if it were brought as a combined effort and identified as such, I think it might have a little more weight.

MEMBER HALPERN: I would concur with that, speaking as a vascular surgeon. I think we all do need to get on the same page because it is better for our patients.

CO-CHAIR MORRIS: Okay. Let's move on to the next set, which is general, ophthalmology, orthopedics, and pediatrics. And we are going to go slightly out of order here. We will start with 0352, Dr. Siperstein, failure to rescue in-hospital mortality, risk-adjusted.

CO-CHAIR TORCHIANA: If I could

just interject, we haven't heard from the developer on this one as yet.

CO-CHAIR MORRIS: Hi, Jeff. Do you want to speak to the group?

MR. SILBER: I am sorry, was I supposed to give a brief overview, failure to rescue? Is that --

CO-CHAIR MORRIS: Only if you want to.

(Laughter.)

We are inviting you to give a brief overview of the CHOP measures.

MR. SILBER: Okay. If there are no specific questions, this is a measure that has been previously endorsed, and this is in the renewal process. We have updated codes for both the 30-day measure and the in-hospital measure.

Failure to rescue is an outcome measure that is risk-adjusted. It is something that complements mortality and complications statistics and has been used for

a number of years to examine how hospitals handle patients who develop complications.

Let me leave it at that then.

CO-CHAIR MORRIS: Okay.

MEMBER SIPERSTEIN: Okay. So, this measure is actually paired with the following measure, and I will explain why. The concept of failure to measure, as was explained, is the probability of death following a complication. And so, the ratio really looks at the number of deaths divided by the number of patients who have had complications.

And this concept is somewhat interesting in the quality arena in that the concept is that just looking at crude mortality, even if risk-adjusted, may not sort everything out for you.

It also is paired with the concept that various safety measures that we do act to reduce, but really do not completely eliminate complications.

And the other concept is that a good number of complications may be due to underlying patient conditions, things that we may not have control over.

And so, the concept is that once a patient has had a complication in the hospital, are the systems in place to prevent that from progressing to death? That is the bottom-line concept.

So, it is somewhat of a different dimension of quality than we classically look at. This measure proposes to look at general surgery, orthopedic, and vascular surgery cases, a large bundle of CPT codes for each of those.

Obviously, the feeling is that it is an important measure. There have been lots of papers written on it.

It garners administrative data. So, it is going to be fairly easy to gather the information.

The measure is a maintenance

measure, but this particular measure has not been used yet in public reporting.

The difference between this Measure 0352 and the next one, 0353, differs only in that this measure looks at in-hospital mortality and the following measure looks at 30-day mortality.

And in our conference call, we could clearly understand through administrative data how you can get the in-hospital and had some questions in terms of the method of collection of the 30-day.

Also, I would parenthetically add that there is another measure that we are going to be reporting that is very similar to this that has been in use and is used by a number of states already.

MR. SILBER: Yes, I was going to mention that we envisioned that the 30-day measure could be used by datasets where 30 days is available, like Medicare data; whereas, HCUP data would only have in-hospital

data. So, that was one of the reasons to have both a 30-day measure and an in-hospital measure.

Also, the other measure, the AHRQ measure, is very, very different. That is a spin on failure to rescue, but it only includes patients who had a subset of the complications. So, that they basically use about half the deaths that would be available in my original measures. So, they are very, very different measures, and I have written a paper to show that, which is referenced in the packet.

CO-CHAIR MORRIS: So, Jeff, does that mean that the 30-day mortality measure should exclude patients who aren't in Medicare or some other system that catches 30-day mortality?

MR. SILBER: If you are using the 30-day measure, then you need a dataset that has 30-day mortality, such that, for example, Medicare has the ability, the beauty of

Medicare data is that it has the ability to look at death at 30 days because you have the denominator falling; you know when patients die.

The problem with HCUP data is that it is in-hospital data, not linked over time. So, there are many situations where you can't do 30-day mortality, and so, you are left with in-hospital mortality.

There are also situations where, for whatever reason, one wants to use in-hospital mortality. Most of the time, we will report our results using 30-day, if we have the ability to report 30-day data.

So, that is why I think it is important to have both measures, because there are many studies that are done or many analyses that are done with data that isn't linkable, and therefore, they should have a uniform way and a reasonable way to compute failure to rescue for in-hospital. And if it is linkable, then they have the 30-day

measure.

CO-CHAIR MORRIS: For the patients in whom you are measuring 30-day mortality, should there be an exclusion to not include patients that are not in Medicare?

MR. SILBER: No. It depends on the dataset that you have. If you have data that is linkable, it is not just Medicare data; there are other datasets. I gave that as an example. Maybe I am not understanding your question, but --

CO-CHAIR MORRIS: Okay. So, what we are talking about here is in reporting that hospitals or facilities could choose whichever dataset they have and report one measure or the other?

MR. SILBER: Well, just like someone would report 30-day mortality or they would report in-hospital mortality, this is comparable, 30-day failure to rescue or in-hospital failure to rescue.

So, when I think about, when I

look at different reporting sets, for example, I mean there are state reporting systems where they are not linked to death certificates, and, therefore, they often use in-hospital mortality. There are other reporting systems where they do have linkable data, and they can look at 30-day mortality.

The original reason to use 30-day mortality was that you didn't want to miss, you didn't want to have the problem of being discharged sicker and quicker and dying outside the hospital. That was always the worry, the bias that might occur if there were differences in length of stay that could lead to different biases, lead to biases in the reporting of hospitals that discharged patients very quickly. Maybe their death rates would look better.

If you have a linkable dataset and you can use 30-day mortality, you don't have to worry about that particular bias. That has always been the drawback with in-hospital

mortality, is that people have been worried that hospitals with different styles of practice might show up differently in terms of in-hospital.

So, we think it is important to have both measures. There are problems when people are more concerned; there are situations where people are more concerned about that, and they should be able to do 30-day, or if they don't have 30-day data, use in-hospital.

Just like people will report both rates, 30-day and in-hospital, for mortality, we would think it would be important to be able to report both rates for failure to rescue.

CO-CHAIR MORRIS: Okay. Thank you.

MEMBER DUTTON: Given that you are risk-adjusting the results, why are you cutting it off at 90?

MR. SILBER: Oh, well, there was a

situation where we were worried that one might have a greater chance of not wanting to rescue after some age. So, we decided that 90 was a reasonable cutoff. In other words, that the intent to rescue might be different when one gets exceedingly old.

MEMBER DUTTON: I think that is ageist.

(Laughter.)

But, aside from that, if it is in your risk-adjustment model, it shouldn't matter. And why not get all the data and report all the data?

MR. SILBER: Yes, we have gone back and forth on this over the years and came out with 90 being what we thought was a reasonable way to look at the numbers. I can't imagine there would be huge differences between the two.

And our sense is that, you know, whenever you are using a mortality measure, there are always issues with mortality that

relate to questions of DNR, et cetera, which generally in these kinds of databases you don't have uniform DNR reported. You don't have it collected or reported uniformly. So, that is why we decided to use that 90 cutoff.

I would think that if one was using public reporting and it included patients up through 90, that that would still do a very good job at getting at whether the hospital has the ability to handle patients who develop complications.

And in terms of being an agist, which I am not, you know, we looked at this mainly from a measurement perspective and worrying about unobservables. So, you know, I think that if there was a concern that some hospitals were specifically undercaring for the very, very old, a change in that definition could be made. But I would say that in most of the work that we have done, it made sense to have some upper bound. That is what we have developed up until now, and that

is the way we have developed the measure, not thinking that it would have a huge change, but we thought it would be slightly less -- put it this way: the data would be a little bit more comparable if we kept it at that upper bound of 90. But I would have to consider going back and seeing if there was some great need to change that.

MEMBER WILHOIT: One question I have is the measure is looking at the percentage of patients with a complication who died. But in the numerator and the denominator, it looks like people who didn't have a complication who died are also included. And I am sort of trying to understand how you describe the measure, then, if you are including people who didn't have a complication, but it is about people who had complications. That is what I am having trouble with.

MR. SILBER: Right. And this is the way that I have developed this since the

very beginning.

I think it is important to look at this in terms of why people die. And, in general, the general way that I have looked at this is that, in order to die or the probability of death is equal to the probability of dying, given a complication, times the probability of a complication, plus the probability of dying, given no complication, times the probability of no complication.

What we have thought was that, since the probability of death, given no complication, is generally pretty low in the surgical arena, that what we are really looking at is that death equals the probability of a complication times the probability of death, given a complication.

There are instances where people die without complications. We call that, we call the opposite of that the precedence rate. We want to develop a measure that covers the

death. So, the way we construct the complication list is that we try to get most of the patients, the great majority of patients who died, we try to have a list of complications that includes the patients who died.

So, for in-hospital, that means about 95 percent of the patients who died had a preceding complication. In the cases where patients die without a preceding complication, the assumption is, if they went through surgery, they probably had something go wrong with them before that.

So, we make sure that we cover all the deaths. And so, the failure rate is, it is always an adjusted measure, but the numerator are the number of deaths; the denominator are the number of complications plus this very small, little piece, which is the number of patients who died without a complication but still underwent surgery.

So, the number, doing it that way

allows us to count all the deaths. And what we basically say is that this is an undocumented complication, that the patient had an undocumented complication. They had surgery. They died after surgery. Something went wrong. In some way, we can document that. In some way, we know that something didn't go right and the patient died.

So, if we are talking about these conditions that we have been studying, you know, the situation is that these patients had something that went wrong. Otherwise, they wouldn't have died. They weren't dead when they went into surgery. So, that is why it is kind of filling in that gap.

The very, very different thing is used with the AHRQ measure, where they use a subset of complications, but they never worry about covering all the deaths. And therefore, they lose about 50 percent of the deaths with a very, very limited set of complications.

The way I set up my list of

complications is I start by saying, what are the complications that are associated with death, and try to get all those complications. When you do that, you end up with in the elderly population, roughly speaking, between 40 and 50 percent of patients have some complication.

If I would not be concerned with covering the complications, but only use deaths following the complication, as they do in the AHRQ definition, then I would end up with greatly disturbing the reliability of the measure because the number of deaths gets cut in half.

CO-CHAIR MORRIS: Okay.

MR. SILBER: So, there is that small, little piece of patients who die without a complication. It is very rare, and it has to do --

CO-CHAIR MORRIS: Yes, you said that.

MR. SILBER: -- with these

differences in the way the claims data pick up some of the complications. But that is a small piece that we try to minimize by setting up a broad-enough list of complications, which is what we have presented, that cover almost all the deaths.

CO-CHAIR MORRIS: All right. Thank you.

I think that that has actually very good face validity. Whenever there is a death within 30 days after an operation, we have to assume that the operation is what led to the death. Even if that is not actually true, that is the assumption, so that we can take the best possible care of our patients and scrutinize our practices to try to make them better. So, I think that seems very valid to me surgically.

Anybody else in the room want to speak to that?

MEMBER DUTTON: Well, one exception might be trauma, where the patient

comes in and gets a desperate procedure in a bid to save their life and then dies. That is not included in here.

CO-CHAIR MORRIS: And it still gets presented in the morbidity/mortality conference.

MEMBER SIPERSTEIN: Yes, but this covers general surgery, orthopedics, and vascular.

MEMBER CARPENTER: I think these are innovative, interesting, and seem to be valid measures. The trouble is in the public understanding of these and the public reporting because they are complicated, and they are a little harder to understand.

Now that doesn't mean we don't endorse them or think about that, but I think that creates some degree of difficulty using these measures as publicly-reported measures. If everybody uses them and we can compare institution-to-institution, then that maybe makes it easier, but they are a little hard to

wrap your brain around, unless you have been doing it for a while.

MEMBER SIPERSTEIN: Actually, if you don't understand anything about it, it is simple understand because a lower death rate is better.

(Laughter.)

If you understand a little, it becomes complicated. And if you understand everything, it becomes simple again.

(Laughter.)

MEMBER DILLON: One of the problems in interpretation of the 30-day result, the mortality, is going to be there is now increasing data that shows that the ultimate disposition of the patient in care outside of the initial hospitalization or institution has a significant impact on readmission rates, subsequent postoperative complications. This is some information that has come out recently in the surgical literature.

So, how does your measure take into account those variables that are outside the control of the initial hospitalization, if you will, and those involved?

MR. SILBER: Right. Well, I mean, looking at mortality, it would have the same issue. You know, if you looked at 30-day mortality, you would say, well, it is possible that it wasn't our fault, that something went wrong outside the hospital, and therefore, the patient died.

It is always going to be the case when you have a measure that doesn't include just the in-hospital, but at the same time we all recognize that, in part, the hospital is in part responsible for this in terms of where they discharge the patient to.

So, you know, I would say that we wouldn't, as a group, say no one can look at 30-day mortality. It is often the gold standard. But it does suffer from that problem that you mentioned, and it has to be

thought about in the same way failure to rescue would be suffering from that same issue.

Anytime you talk about a mortality rate that is occurring outside the hospital, our sense is that, you know, to get a good picture of the hospital, you are still going to need 30-day, that most of the time most of the deaths are occurring fairly early.

Secondly, the alternative of only looking at in-hospital has its problems and its bias, too.

So, what you are saying isn't unique to failure to rescue. It is true for any measure that has a set point in time that can occur outside the hospitalization, but those are often considered to be the gold standard measures.

And if there are caveats about who is to blame for this, I mean I think that has to be addressed with future research, and there might be certain situations where people

are very, very worried about this. But if you are evaluating a hospital, they, in part, should be responsible for, when they discharge the patient, what happens to them.

The same is true for readmission, right? And maybe are we going to be more --

CO-CHAIR MORRIS: We are not going to talk about readmission right now.

MR. SILBER: Right. Okay.

CO-CHAIR MORRIS: Anybody else on the Committee want to make any comments or ask any questions?

MEMBER STAFFORD: I just have one more question, and maybe it is philosophical. But I would like to know why a group from CHOP would propose an adult measure, and why you wouldn't propose a pediatric measure. I understand the complications in that population are different, but why? It would seem to me that your vested interest would be in your pediatric population.

MR. SILBER: Well, I received my

Ph.D. in health services research in 1990 using Medicare data, and I have been doing research in both adult and pediatric diseases since that time. So, I have a long experience with using claims data in adults and thinking about surgical problems in adults and pediatrics. So, I have never let the particular age influence my research, age of the patient influence my research areas.

And the nice thing about being a CHOP is, you know, I am a professor at University of Pennsylvania, I am a professor at Wharton, I am a professor at the Medical School. They have been wonderful about letting me do my work in all different areas.

The fact that my primary appointment is in pediatrics is why this is a CHOP measure, why the word Children's Hospital of Philadelphia is there, but, you know, I have been doing work in adult surgery for over 20 years.

So, that is why it happens to be

just the fact that I originally am Board-certified in pediatrics and I happen to be at Penn in the Department of Pediatrics. I am also in the Department of Anesthesiology and Critical Care, both adults and pediatrics.

CO-CHAIR MORRIS: Okay. Thank you very much.

Anybody else on the Committee have any questions specifically for Dr. Silber about this measure?

MEMBER ROGERS: I have a quick question for Carol. Do you have any concern that the basis for this is administrative data, recognizing the little dissonance between what actually happened and what gets coded?

MEMBER WILHOIT: I think I don't have a problem with this. And I think that we have used the AHRQ measure for a number of years, which, again, has some differences. But, I mean, administrative data is useful. There is just some things correlate better

than others.

MEMBER DILLON: Is it going to be good enough to pick up conditions present on admission?

MEMBER WILHOIT: It doesn't look like this has present on admission as a consideration. So, that obviously would be a question, you know, as to whether it should.

MEMBER DUTTON: This looks like it is based on the presence of an ICD-9 complication code. And I was wondering if there was an increasing incentive to not report those, if you don't get paid for that admission.

MEMBER WILHOIT: And actually, we are seeing more and more codes being submitted rather than less and less, for a variety of reasons, probably because of present on admission, and so on.

CO-CHAIR MORRIS: Okay. Are we ready to vote on Measure 0352, failure to rescue in-hospital mortality, risk-adjusted?

First of all, does the measure meet NQF criteria for importance to measure and report?

(Vote.)

And the results are 18, yes; three, no.

Does the measure meet NQF criteria for scientific acceptability of measure properties?

(Vote.)

Nine say completely; eleven, partially; one, minimally.

Does the measure meet NQF criteria for usability?

(Vote.)

Seven say completely; twelve, partially; two, minimally.

Does the measure meet NQF criteria for feasibility?

(Vote.)

Will that 21st person hit your button again and, then, hit Send?

Eight say completely; twelve, partially; one, minimally.

And, then, lastly, does the measure meet all of the NQF criteria for endorsement?

(Vote.)

Eighteen say yes; three say no.

The next measure is --

MEMBER DILLON: Can I just one quick question?

CO-CHAIR MORRIS: Sure.

MEMBER DILLON: So, one of the statements made at the beginning of this was that this was not publicly reported, if I am not mistaken. Did I hear that correctly, that this is not in the public domain in terms of reporting?

MEMBER SIPERSTEIN: No, it is just not currently in use.

MEMBER DILLON: Okay. Does the fact that we have endorsed it, will that automatically move it into public reporting or

is there a whole other series of steps --

MEMBER SIPERSTEIN: No, my understanding is that it is a maintenance measure. It just has not yet been picked up and adopted by any organizations.

MS. MURPHY: So, what you would want to hear is the plan for public reporting within this endorsement period, within this next three years.

MEMBER STAFFORD: So, before we go into the next one, can I just make a comment and make sure that we somehow at some point, when we are talking about harmonization, harmonize these with the already two endorsed NQF measures that are AHRQ measures? So, that is 0220 and 0351. They were both, it looks like, last endorsed September 2010, according to the NQF website.

MS. MURPHY: So, the consideration of harmonization or competing measures will be taken up around these three measures.

MEMBER STAFFORD: Okay. Great.

CO-CHAIR MORRIS: Okay. The next one is 0353, failure to rescue 30-day mortality.

We had a little bit of a discussion about, well, we really discussed this already. And we can move on to the vote.

The one thing that still isn't truly clear to me is whether hospitals are supposed to go after 30-day mortality among non-Medicare patients.

It is nice to have alternative means of measuring these things, but the way that it stands right now, it looks like this is not really an alternative measure to the first one, that it is an additional measure to the first one.

So, for non-Medicare patients, how feasible is identifying 30-day mortality among postoperative patients? Anybody want to say anything else about that before we go ahead and vote?

(No response.)

Okay. Does the measure meet NQF criteria for importance to measure and report?

(Vote.)

And please hit your button again and hit Send again.

MS. MURPHY: He is out of the room.

CO-CHAIR MORRIS: We are missing somebody?

MS. MURPHY: We only have 20.

CO-CHAIR MORRIS: Okay. Seventeen say yes; three say no.

Does the measure meet NQF criteria for scientific acceptability of measure properties?

(Vote.)

Six say completely; twelve, partially; two say minimally.

Does the measure meet NQF criteria for usability?

(Vote.)

Three say completely; ten say

partially; eight say minimally.

Does the measure meet NQF criteria for feasibility?

(Vote.)

Three say completely; ten say partially; seven say minimally, and one says not at all.

And lastly, does the measure meet all of the NQF criteria for endorsement?

(Vote.)

Thirteen say yes; eight say no.

So, among our considerations for this, one is that we would like to know the plan for public reporting, and the other is that we would like to look at harmonization among the other measures and the AHRQ measures.

MR. SILBER: Thank you.

CO-CHAIR MORRIS: Okay. Great.

So, we are going to actually jump around a little bit here and go to 1550.

MR. SILBER: I just want to say

thank you very much. And I guess we are talking with you again tomorrow? Tomorrow?

CO-CHAIR MORRIS: That's right. Yes.

All right. Thanks a lot, Dr. Silber.

MR. SILBER: Bye-bye.

CO-CHAIR MORRIS: So, 1550, hospital-level risk-standardized complication rate following elective primary total hip arthroplasty and total knee arthroplasty. And this is being presented by Dr. Carpenter.

MEMBER CARPENTER: Yes, I will present this. This is a newly-proposed measure.

CO-CHAIR MORRIS: I'm sorry, let me interrupt you for just a second.

So, another developer is going to speak?

DR. BURSTIN: The developer haven't had a chance to do their brief update, if that is okay.

CO-CHAIR MORRIS: Okay.

DR. BURSTIN: She was on the train earlier, if you recall.

CO-CHAIR MORRIS: Got you. Okay.

Would you like to give a brief update?

MS. GROSSO: Yes.

CO-CHAIR MORRIS: Into the microphone.

MS. GROSSO: So, we developed two complementary administrative claims-based measures for patients undergoing primary elective total hip arthroplasty and total knee arthroplasty.

We developed a risk-standardized complications measure and a 30-day all-cause readmission rate.

We developed the measures for these procedures because they are high-volume, high-cost procedures. In a cohort of Medicare fee-for-service patients in 2008, there were over 100,000 hip arthroplasties performed

across 3,000 hospitals and 241,000 knee arthroplasties performed across 3300 hospitals. And that was after we excluded patients with hip fractures and patients undergoing partial hip arthroplasties, revision procedures, and resurfacing procedures.

And the trend in performance of these procedures is projected to increase. The cost is high, and that is projected to increase also.

We developed these measures over 12 months with extensive input from nationally-recognized leaders in the orthopedic and surgery community, including the past and current Presidents of the American Academy of Orthopedic Surgeons, the past President of the American Association of Hip and Knee Surgeons.

And again, we developed two measures that assess separate domains of quality, a complications measure that captures

overall care and reflects the technical aspects of the procedure. And that is targeted toward, you know, geared toward more targeted QI improvement efforts. And a readmission measure which globally measures care, including transitional care.

It is a risk-adjusted measure. It is risk-adjusted for patient case mix, age, gender, number and type of procedure.

Thank you.

CO-CHAIR MORRIS: Thank you.

Dr. Carpenter?

MEMBER CARPENTER: Yes. Thanks.

So you get a flavor for what these are, these are useful to think about together.

The complications data in terms of importance, the impact, again, they are high-volume, costly procedures and they are growing. The overall rate of these complications is about 7 percent. So, it is not super-high, but it is significant. The range is around 2 to 9 percent that they

report. So, there is quite a variable range from different institutions.

They did limit the complications to what I think are appropriate ones, including death, mechanical complications, infection, bleeding, PE, MI, pneumonia, and sepsis. So, it is not every little, tiny complication. There can be some debate about how severe a wound infection is, but other than that, I think it is a good option.

The followup timing is interesting because it is 90 days for some of these complications; it is 30 days for others, and seven days for others. So, that makes it a little bit more complicated.

But it is a claims data measure. So, hopefully, that is all available.

It is limited to patients 65 and older, I think because that is where the data is good. A lot of patients, I don't know the percentage, but I would say 25 or so percent, maybe 30 percent of these procedures are done

in patients under 65. So, there is a segment of people that are not included in these measures.

And exclusions I think are appropriate, transfers, hip fractures, and revision surgery.

The risk-adjustment method is complicated, sophisticated. I don't think I understood it all, but it is a hierarchical linear regression that ends up reporting a predicted complication rate versus an expected complication rate, predicted based on your past performance of your institution and how you would do, and expected based on the mix of patients that you have. So, it is sort of like an observed to expected. That is how I understood that.

So, I think it seemed valid. There is a lot of math in there. So, it is a little hard to understand, but I think it is valid.

And it is similar to other

measures that I think are already approved for MI and heart failure. So, some of those same techniques have been used, I assume successfully, although I don't know.

So, that is basically the complications one, and maybe we should just stop there. Then, we can discuss the other one, but once we get through this, the other one won't take very long.

And just in terms of the -- I was not on the conference call. So, I don't know all the discussion. There was some concern about the 65-year-olds, why it was limited to that, and some discussion about the risk-adjustment technique.

CO-CHAIR MORRIS: Can somebody that was on the call speak to that, the concern about the risk-adjustment technique?

MEMBER AFSAR-MANESH: I don't remember what type of discussion we had exactly. I know we spoke about the different number of days for each of the complications.

And part of the thought was that likely there is something in the literature that correlates seven days for sepsis. Thirty days or 90 days is not going to be able to be tracked back to that particular surgery. But that is the only discussion I remember having.

MEMBER DILLON: And I would agree. I think it was predominantly on the variability in the length of followup for the complications.

And, then, just two questions or one question about two outcomes. Readmissions and re-ops, were they factors that were looked into and that dropped out of the analysis? Or have they not been included?

MEMBER CARPENTER: So, this measure, the first measure is just those complications I listed. So, they had readmissions that we will talk about next as a separate reportable measure, but, generally, they are going to fall into one of these categories: mechanical failure, such as a

fracture or dislocation, or an infection or bleeding. So, generally, the reasons for reoperation and readmission would be captured with these complications, if they happened in that timeframe.

CO-CHAIR MORRIS: Can you clarify why acute MI, pneumonia, and sepsis were not included in the 30-day window, but rather in a seven-day window?

MS. GROSSO: Yes. It was clinically-driven, and we wanted to capture the important medical complications. And we felt that in order to be attributable to the procedure, we set the cutoff at seven days because anything beyond that, it would be hard to attribute in this population to the procedure.

CO-CHAIR MORRIS: Jim, what is your perspective on that?

MEMBER CARPENTER: Well, I think they had a lot of discussion. I think I would applaud them for including a good group of

people experienced with these procedures in their Technical Panel. So, I think that was probably well-vetted and gone over pretty thoroughly. It seems reasonable to me.

MEMBER DILLON: It is a philosophical problem, though, because, again, if the public perceives that most surgical outcomes are going to be framed within a 30-day followup period, and now they have to -- well, wait a minute. It says seven days; it says 14 days.

To me, one of the things that we should be discussing is somewhat of a standardization of reporting these outcomes. And to have different ones for different procedures in different areas concerns me.

DR. BURSTIN: I just had a question of whether the Yale team had actually done any analyses to see if you maintained all of them at 30 days, which would be certainly more standard, how much higher did the complication rate go, and was there

significant attribution? I mean, in general, we are kind of going away from the idea that it has to be directly related to the procedure, as you do for all of the all-cause readmission measures. So, I am just curious.

MS. GROSSO: We didn't look at that specifically. But what we did do, we looked at the trend in the rates over 90 days. And for the AMI, pneumonia, and sepsis, they peaked within seven days and then dropped and went down to baseline.

And we looked at the trend over 90 days for all the complications. So, that was clinical input, and based on an analysis of the data, that informed the followup, the decision to set the followup period for each complication.

DR. BURSTIN: If a patient goes to a SNIF, are they included still? Just a question.

MS. GROSSO: If they are discharged to a SNIF? Yes. Yes.

MS. BERNHEIM: And just remember that part of the response to your question about the rehospitalization is that this is paired with a readmission measure. So, those sort of readmissions that are less related to the procedure, but more about the sort of overall quality of care and transition are going to be captured in the complementary measure. So, that was part of the thinking there as well.

CO-CHAIR MORRIS: Any other discussion, comments?

MEMBER STAFFORD: Yes, I just have a quick question. And you talked about it in your exclusions, but it still isn't clear to me why you would exclude DVT and UTI, just because screening practices are different.

MS. GROSSO: Yes, we did have a lot of discussion about that. The codes that are used vary, and particularly the DVT. Screening varies considerably across sites, and readmission for screening varies.

The complications we chose are associated with readmission. You had to get readmitted for the complication.

And the surgeons felt that, with the other two, it would just be some of the codes aren't very specific, particularly those for UTI, and with the DVT there was just so much variability in screening and readmission rates that it wouldn't be helpful to include.

MEMBER STAFFORD: But DVT is one of the more common complications, and people don't always screen for it, you're correct. It is actually often picked up when somebody complains about it. So, routine screening for DVT in most surgical populations isn't done. But it would seem to me that you are going to miss a significant number of patients.

I am not an orthopedist, but that is one of the things that the orthopedists all worried about, and this is what we all argue about when it comes to what kind of DVT prophylaxis you use.

So, certainly, at least DVT I think is something that you really need to think about, and UTIs in the elderly population are something that are common enough as well.

And if you don't throw those out -- if you throw those out, then maybe your teams aren't getting those Foley catheters out within 24 hours, which is one of the SCIP measures. So, just a couple of thoughts.

MEMBER CARPENTER: Well, I will make a comment about the DVT, I think. Obviously, they discussed this thoroughly, but the problem, this is a claims database. So, if you could identify symptomatic DVTs versus ones picked up that are asymptomatic from screening, that might be helpful. But because this database couldn't separate those out, it is really going to be dependent on how much screening is done in those patients because the risk for a below-knee asymptomatic DVT in these procedures is very high.

So, if you are going to screen everyone, you are going to find a lot of those. If you are going to screen no one, you are not going to find very many of them, and especially in this window.

So, I think it is not that it is not important. It just probably wasn't reliably in the measure enough to include. And probably a measure isolated just on DVT prophylaxis and measurement should be developed, but right now I don't think the technology and the use is there.

MEMBER WILHOIT: It was mentioned that one of the concerns the Work Group had raised was the age 65. Can you explain why it isn't including younger people?

MS. BERNHEIM: Sure. This is really at this point a matter of data availability. As with our other measures, we have developed this in the Medicare claims data, and within the Medicare claims data the younger-than-65 group is a very narrow group.

To get into Medicare before 65, it is, yes, end-stage renal disease patients and patients with disabilities.

We have been doing a lot of work to look at our measures and see how well they work in broader populations. And like this, as with this measure, we actually expect that the measure would work pretty well in those are the populations and would hope to have its use in those populations eventually. But, at this point, it was just matter of data availability.

DR. BURSTIN: And we have had some discussions with CMS about bringing to us all population-level measures. So, I guess since it has already been brought up here as well, again, one recommendation from the Committee could be to specifically request that those specifications, when available, come to NQF because, again, this comes up again and again, and then we wind up getting measures for the below-65 population and we spend inordinate

months harmonizing them. So, we try to avoid that.

MS. BERNHEIM: That work is underway.

CO-CHAIR MORRIS: Any other questions or comments?

(No response.)

Okay. Let's go ahead and vote.

On Measure 1550, does the measure meet NQF criteria for importance to measure and report?

(Vote.)

Nineteen say yes; one says no.

Does the measure meet NQF criteria for scientific acceptability of measure properties?

(Vote.)

And eleven say completely; eight say partially, and one says minimally.

Does the measure meet NQF criteria for usability?

(Vote.)

Ten say completely; ten say partially.

Does the measure meet NQF criteria for feasibility?

(Vote.)

Fourteen say completely; ten say partially. Sorry, six. I don't know where that came from. Oh-oh.

(Laughter.)

Does the measure meet all the NQF criteria for endorsement?

(Vote.)

Twenty say yes; zero say no or abstain.

All right. And, then, the next measure, 1551, hospital-level 30-day all-cause risk-standardized readmission rate following elective primary total hip arthroplasty and total knee arthroplasty.

We have really pretty much had the discussion here, but, Jim, did you have more to add?

MEMBER CARPENTER: Not really. It is actually a simpler measure than the previous one. And I think it is because they use all causes for readmission other than readmission for elective procedures, such as the other hip or the other knee.

I think it is simpler. They use the same risk-adjustment strategy, and it is all 30 days. So, again, that makes it simpler. So, I think this is a little more straightforward than the other.

CO-CHAIR MORRIS: Any other comments from the group or questions?

(No response.)

Let's go ahead and vote.

Does the measure meet NQF criteria for importance to measure and report?

(Vote.)

Twenty say yes.

Does the measure meet NQF criteria for scientific acceptability of measure properties?

(Vote.)

And fifteen say completely; five say partially.

Does the measure meet NQF criteria for usability?

(Vote.)

Sixteen say completely; four say partially.

Does the measure meet NQF criteria for feasibility?

(Vote.)

Fourteen say completely; six say partially.

And lastly, does the measure meet all the NQF criteria for endorsement?

(Vote.)

Nineteen say yes; one says no.

DR. BURSTIN: Just one clarification. Does the developer intend these to be truly paired measures, meaning they should always be reported together? You used the word "paired" earlier. I just want

to check if that is what you intend.

MS. BERNHEIM: They were developed to be complementary, but we did not intend that they couldn't be separated. Yes. Right.

DR. BURSTIN: That means something very different; that's all, I guess.

MS. BERNHEIM: Okay.

CO-CHAIR MORRIS: Okay. Since we had our discussion not that long ago around the topics of 0351, we will go back to that, 0351, death among surgical inpatients with serious, treatable complications. This is the AHRQ measure.

MR. ROMANO: Could I provide a little background on that from the developer's perspective?

CO-CHAIR MORRIS: Yes. Thank you.

MR. ROMANO: Right. So, I appreciate hearing the discussion of the earlier indicators. So, let me provide a little bit of historical perspective that may not have been clear from the documents.

So, this measure was definitely inspired by the work that Jeff Silber and his colleagues have done over the last 20 years. So, it reflects the same underlying concept of what he calls failure to rescue or what we call death following serious, potentially-treatable complications.

At the time that this measure was developed by AHRQ as part of the Quality Indicators Program, we surveyed the field to find different specifications of this concept for administrative data, because the original specification was actually based on a more complex dataset that was a proprietary dataset.

Jack Needleman and Peter Buerhaus did a large study that was funded by HRSA looking at the impact of nurse staffing. And they convened an expert panel to help them identify measures that would be sensitive to nurse staffing and nursing skill mix.

So, they identified this as such a

measure, but they specified it based on patients having one of six complications to get into the denominator. So, their expert panel felt that these six complications were complications that were particularly serious and where a rapid response effort could identify the complication early and potentially slow or prevent the progression.

So, Needleman and Buerhaus operationalized this specification based on these six complications for the denominator, showed that it was valid based on their work on their staffing.

We were familiar with that work and consulted with Needleman and Buerhaus. We convened our own expert panel under AHRQ's auspices, and they basically agreed that they liked the concept of failure to rescue, but they preferred a specification in which the denominator was linked to specific types of complications that they could get their arms around clinically, where they felt that there

was a particular opportunity for early intervention and, thus, actionability for improvement.

So, that was the concept, and that is how this measure came to be the original NQF-endorsed measure of failure to rescue.

Jeff Silber and I are friends. I was actually a co-author on the paper that he described comparing these different specifications. So, I have nothing bad to say about his measure or his work.

But he decided to seek NQF endorsement for his measures separately. And at that time, the AHRQ measure was already endorsed, and there was no mechanism at that time for harmonization. So, this is an opportunity for kind of revisiting the underlying question and potentially harmonizing.

We did, Jeff and I did together a comparative analysis of these specifications. And as you might expect, his specification

offers somewhat higher reliability because of the fact that it includes all the deaths as opposed to only including half the deaths.

The construct validity is actually similar in terms of the correlation with measures of nurse staffing, skill mix, teaching status, Board certification of physicians, and so forth. So, measures of construct validity are similar.

The AHRQ measure is a little somewhat less sensitive to patient characteristics. The Silber measure is somewhat more sensitive to patient characteristics, and, thus, potentially to confounding due to those characteristics.

Those numbers are all on the paper. I am happy to share that with the group, if you are interested.

So, in terms of going forward, obviously, the AHRQ measure is available in public-use software, and therefore, it has been picked up more broadly in public

reporting applications by a variety of states and other reporting organizations. That doesn't necessarily mean that it is superior. It is just the product of a different expert panel process with Needleman and Buerhaus' panel on nurse staff and, then, our panel on the inpatient quality indicators.

So, we are open to your suggestions and comments related to opportunities for harmonization.

CO-CHAIR MORRIS: Thank you.

Dennis?

MEMBER RIVENBURGH: After that, there's not tons to say. Again, this is a maintenance measure that looks at outcomes data based on all discharges with a disposition of deceased.

And again, unlike 0353, which talks about 30-day mortality, this is only in-hospital mortality that they are going after the data on. So, the dataset is a little bit smaller.

We discussed this in the group, and the consensus really was that these two measures really needed to be brought together. Actually, the three of them needed to be completely paired and kind of brought together into maybe one, definitely two measures that could be harmonized and utilized.

And probably the biggest advantage so far with 0351 is that it is being reported to the public, unlike the other two which are collecting data that is just collecting. So, that is clearly the thing the Committee felt that was the most efficacious of this, is that it is currently being reported.

CO-CHAIR MORRIS: It sounds like there is another potential advantage, which is the goal of this is to improve rescue.

MEMBER RIVENBURGH: Right.

CO-CHAIR MORRIS: And if these are conditions for which there is potentially a more easy-to-identify rescue process, then that also could be advantageous. And maybe

that could be considered during what I hope would be a harmonization process.

Anybody else have questions?

MEMBER WILHOIT: Another advantage is the age, which goes down to 18, I think. But that is really important for state reporting, for health plan reporting, for all kinds of things, to include that broader population.

MEMBER RIVENBURGH: And it actually even goes lower than that with certain conditions.

MEMBER DUTTON: Feel free to include the 90-year-olds.

MEMBER RIVENBURGH: I'm sorry. No, it is 18 is the lower limit.

MEMBER DUTTON: Yes, my grandmother is going to kick you in the shins.

(Laughter.)

CO-CHAIR MORRIS: Anybody else?

(No response.)

Let's go ahead and vote then.

Does the measure meet NQF criteria for importance to measure and report?

(Vote.)

Nineteen say yes; one says no.

Does the measure meet NQF criteria for scientific acceptability of measure properties?

(Vote.)

Thirteen say completely; seven say partially.

Does the measure meet NQF criteria for usability?

(Vote.)

Thirteen say completely; seven say partially.

Does the measure meet NQF criteria for feasibility?

(Vote.)

Fourteen say completely; five say partially.

Does the measure meet all the NQF criteria for endorsement?

(Vote.)

Eighteen say yes; one says no.

So, what we are looking at here is harmonization; also, I think based on the stuff that Richard was bringing up, consideration of changing the upper age limit. And, then, the plan for public reporting among the CHOP measures.

DR. BURSTIN: It is actually more than just harmonization because they are all claims-based measures. I mean it would be difficult, I think, to think about -- we would need to think about how they could coexist or can they come together in a better way?

MEMBER RIVENBURGH: Will there be further discussion of the harmonization question tomorrow?

CO-CHAIR MORRIS: All right. So, moving on to Measure 0339, pediatric heart surgery mortality.

MEMBER DILLON: Do I dare say that we can go through this quickly because we are

dealing with small patients and small numbers?

(Laughter.)

All right, let me just bring this up here.

So, these two measures, certainly one is I think from a no-brainer point of view to discuss and vote on. The 0339 looks at pediatric heart surgery and mortality in a risk-adjusted fashion dealing with congenital heart surgery in an age population 18 and under.

To cut to the evaluation of our group, I think in terms of I think just about all of the categories, our group felt that, though there was some variability in terms of meeting partial to complete concordance with the requirements, that our Work Group felt that this was a valid metric and thought that it met all of the criteria to be proposed and put forward.

There are some concerns in that, in terms of its presentation, it does seem to

lack -- it would be nice if they had included some current data. It is a maintenance measure, and so they have gotten some years of data, and it would be nice to see some of the results.

The problem is, of course, as many of you know, it is limited to a certain number of institutions in the country, and the numbers actually are not increasing. They have been flat to decreasing in terms of overall congenital heart surgery. But our group thought that this was a valid metric for proposal and support.

The volume, if you want me to just continue, 0340 is right on its heels in terms of looking at volume, and clearly meant as a paired process.

Again, in terms of the writeup, the citations are quite old. There is no lack of current data.

Our group, I must admit, felt that for the most part the criteria were met. Many

of us felt that some of it was partially. Certainly, from a scientific and feasibility point of view, the problem I have, again, it gets into a personal bias of dealing with just a pure volume report, particularly in this situation where the numbers are just so small, and whether it is sensitive enough to pick up any differences in terms of institutional performance.

The cut number I think in the literature that was quoted was 200 cases or so, and there are very few programs that are doing those numbers. And as they have cited in the literature, many programs are well below that, 100 cases or less.

And so, again, I am not sure -- this will be, obviously, a personal bias injected here -- that the reporting of volume in pediatric congenital heart surgery is a suitable or a valid measure. Certainly, the mortality is.

CO-CHAIR MORRIS: Anybody have any

questions or comments about that?

(No response.)

Do we want to request from the developer that the references be updated?

MR. ROMANO: Yes, we actually do have updated references. I am sorry if they didn't get to you.

Dr. Welke's group, affiliated with STS, has done some work more recently on this issue. I am not sure if anyone is familiar with that work, basically, showing a consistent volume outcome relationship.

There was one study from California that showed a weakening of the relationship, but the problem there is that the system is already highly regionalized, and one hospital actually had much higher volume than any other hospital in the State. And so, it was a highly-skewed distribution. It wasn't typical of the nation as a whole.

So, more recent literature is consistent, and we will get to harmonization

issues later.

But this measure is really based on the so-called RACs(1) scheme that has been published on fairly extensively in the literature.

MEMBER DILLON: In terms of a volume, the real problem does come in the fact that it is becoming so regionalized. I mean it is getting driven into just certain institutions. And so, for that reason, does volume no longer become a valid measure of quality here?

CO-CHAIR MORRIS: Okay. Let's go ahead and move on to the vote.

I think that is an important point. It probably will become more important over time with other rare operations.

Does the measure meet NQF criteria for importance to measure and report?

(Vote.)

Eighteen say yes; one says no.

Scientific acceptable of measure

properties.

(Vote.)

Thirteen say completely; six say partially.

Usability.

(Vote.)

Fifteen say completely; four say partially.

Feasibility.

(Vote.)

Fifteen say completely; three say partially; one says minimally.

Does the measure meet all the NQF criteria for endorsement?

(Vote.)

Eighteen say yes; one says no.

The next measure, 0340, is up for vote next.

And Dr. Dillon brought up the important point that these are being driven into a few sort of high-volume centers.

Do we want to discuss that now or

is that just something to note?

DR. BURSTIN: Just one point of clarification, that my understanding is the volume measure is paired with the mortality measure. So, the idea would be, if somebody is looking to see which center they want to go to as a parent or someone, that is potentially information you would want to see paired with mortality, though probably not as a standalone, which is why it is paired.

CO-CHAIR TORCHIANA: If I could just say the other obvious thing, the practice of pediatric heart surgery has changed drastically in the last 20 years, probably as much as aortic aneurysm surgery in that sort of the so-called bread-and-butter of pediatric heart surgery, ASDs, BSDs, PEAs, CORs, are virtually never operated on or only occasionally operated on, and almost everything is at a higher level of complexity.

And that influences a whole lot of things, but one of the principal things that

it does is invalidate a big body of literature from the more distant past.

CO-CHAIR MORRIS: Another good reason to update the literature on that.

Let's go ahead and vote.

Does the measure meet NQF criteria for importance to measure and report?

(Vote.)

Fourteen say yes; five say no.

Scientific acceptability of measure properties.

(Vote.)

And ten say completely; eight say partially; one says minimally.

Usability.

(Vote.)

Ten say completely; eight say partially; one says minimally.

Feasibility.

(Vote.)

Thirteen say completely; six say partially.

Does the measure meet all the NQF criteria for endorsement?

(Vote.)

Now everybody please hit your button one more time and hit Send aiming at Jessica.

(Vote.)

Fifteen say yes; four say no.

All right. I would like to kind of take the group's pulse here on Measures 0515 and 0301.

(Laughter.)

Do we want to table these until tomorrow morning?

CO-CHAIR TORCHIANA: Well, do we have people who can't attend?

CO-CHAIR MORRIS: Oh, are you unable to attend tomorrow?

MS. SLOSBURG: We can attend tomorrow, but I was just going to request that it could be in the morning with the other measure because Dr. Shapiro has a commitment.

CO-CHAIR MORRIS: Yes, we will do it first.

MS. MURPHY: Can I just say that updated information about references was sent out to everybody, but it has just been AHRQ updates, within the last day or two. Some of you may also have been traveling. So, among your souvenirs through email there is an attachment with updated references.

CO-CHAIR MORRIS: We are scheduled to start tomorrow at nine o'clock.

In terms of leaving your things in the room, is that not a good idea? We are going to go to another room tomorrow next door, I think, hopefully, a little bit larger. And I think there are more cookies in there.

MS. MURPHY: So, before everyone departs, we do need to open the line for public and member comment.

CO-CHAIR MORRIS: Okay. All right. Any comments?

(No response.)

I don't hear anything. I think we are all set.

Goodnight, everybody.

(Whereupon, at 6:25 p.m., the foregoing matter went off the record.)

**F**

**I**

C E R T I F I C A T E

This is to certify that the foregoing transcript

In the matter of: Surgery Endorsement
Steering Committee

Before: NQF

Date: 05-04-11

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