NATIONAL QUALITY FORUM

+ + + + +

SURGERY ENDORSEMENT MAINTENANCE
2010 STEERING COMMITTEE

+ + + + +

TUESDAY
MARCH 1, 2011

+ + + + +

The Steering Committee met at the Washington Hilton, Lincoln West Room, 1919 Connecticut Avenue, N.W., Washington, D.C., at 9:00 a.m., Arden Morris, Chair, presiding.

PRESENT:
ARDEN MORRIS, Chair, University of Michigan Health system
JAMES CARPENTER, University of Michigan
ROBERT CIMA, Mayor Clinic
CURTIS COLLINS, University of Michigan Health System

PETER DILLON, Penn State Hershey Medical Center
RICHARD DUTTON, Anesthesia Quality Institute
STEVEN FINDLAY, Consumers Union
PAULA GRALING, Inova Fairfax Hospital
VIVIENNE HALPERN, Carl T. Hayden VA Medical Center

EILEEN KENNEDY, Pepco Holdings
RUTH KLEINPELL, Rush University Medical Center
JOHN MORTON, Stanford University
DENNIS RIVENBURGH, St. Anthony's
TERRY ROGERS, The Foundation for Health Care Quality
CHRISTOPHER SAIGAL, UCLA Medical Center

NICHOLAS SEARS, MedAssets
ALLAN SIPERSTEIN, Cleveland Clinic
RENAE STAFFORD, University of North Carolina
CONNIE STEED, Greenville Hospital System
CAROL WILHOIT, Blue Cross-Blue Shield of Illinois
CHRISTINE ZAMBRICKI, American Association of Nurse Anesthetists

NQF STAFF PRESENT:

HELEN BURSTIN
KRISTIN CHANDLER
ALEXIS FORMAN
MELINDA MURPHY
JESSICA WEBER

ALSO PRESENT:
RICHARD PRAGER, The Society of Thoracic Surgeons
DALE BRATZLER, Oklahoma Foundation for Medical Quality
DAVID SHAHIAN, The Society of Thoracic Surgeons (via telephone)

JANE HAN, The Society of Thoracic Surgeons (via telephone)
JESSICA RIEHLE, Ingenix (via telephone)
WANDA JOHNSON, Oklahoma Foundation for Medical Quality (via telephone)

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

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

9:02 a.m.

CHAIR MORRIS: Good morning. Welcome to the second day of the Surgical Quality Measures Steering Committee.

I wanted to just briefly recap some important points from yesterday, and also to once again thank our Steering Committee members for being present and for their effort and attention.

First of all, we need to continue to focus on a couple of things that came out at various times during the day yesterday. One of them is with the maintenance measures, in particular, what have we learned since they were initially endorsed? Have we seen evidence of an impact? Have we learned anything else from the fact that they were enacted earlier?

Secondly, we need to focus a little bit more on the impact on disparities. We are focusing on a lot of important things,

and that is another important thing that we need to focus on.

Thirdly, please be mindful that we are very interested in what the either public reporting plan is or ensuring that public reporting is actually already in existent.

Then fourth, we need to continue to speak to the cost and burden on hospitals, especially for the measures that are associated with proprietary databases, and I think that that came up yesterday several times, and it is important to remain mindful of it.

We would like to give our developers a few moments to introduce the candidate measures that they have for today, and I see that Dr. Prager is here from STS. The first two measures are yours. Would you like to start?

DR. PRAGER: I am happy to start these three measures that are, I think, three measures for today for the STS. Jane Han

introduced the concept yesterday on the phone of these measures and when they started and, essentially, I would presume, we will discuss them in the same format, anti-lipids, anti-platelet agents at discharge, and post-operative deep wound infections.

CHAIR MORRIS: Thanks. Do we have Ingenix on the telephone? All right. So in that case, the third measure is developed by Ingenix, and we will just -- When they get on the phone, we will just have them start, but we may skip to 0130, if we don't have the phone on when it is time to talk about their measure.

First of all, measure -- oh, I'm sorry. Is CMS here? Would you like to introduce your measures as well?

DR. BRATZLER: I will make it really clear. I am not CMS. My name is Dale Bratzler. I am with the Oklahoma Foundation for Medical Quality, and we are a contractor to CMS supporting the hospital inpatient core

measures.

We have three measures that are being considered for reendorsement today. All three are currently in use, publicly reported, and I believe all three, or at least two of them, are a part of the proposed value based purchasing measures for 2013, Fiscal Year 2013.

The first one is cardiac patients with controlled postoperative serum glucose, again a measure limited to cardiac surgery patients, so affects about 1100 hospitals in the United States currently; and then two measures on VTE prophylaxis.

The first one, recommended VTE prophylaxis for surgical patients, and the second one patients who receive appropriate prophylaxis and received it in the appropriate time frame, within 24 hours before or after the end of surgery. Approximately 3500 hospitals currently capture data on those two VTE measures. When the discussion happens, I

am happy to answer any questions.

CHAIR MORRIS: Thank you. The first measure is measure 0116. Dr. Kleinpell? All right, 0116, Dr. Kleinpell, anti-platelet medication at discharge.

DR. KLEINPELL: Sure. The measure number 0116, the measure title: Anti-platelet medication at discharge. The measure steward is Society for Thoracic Surgeons.

The description of this maintenance measure is percent of patients age 18 years and older undergoing isolated CABG who were discharged on anti-platelet medication.

This is submitted for maintenance review. It was first released in 2004, last revised in 2010, and it is indicated it is updated annually.

In terms of importance, we know that the use of anti-platelet therapy at discharge is currently an accepted standard of care to improve bypass graft patency, as well

as promote secondary prevention of coronary artery disease.

So the measure is important. It is also currently a CMS PQRI initiative. It is 169. The information that was provided to us was that there still is a performance gap. Despite the fact that it has been around for a while, the information noted in a sample of 581 patients was that the performance ranged from 85 percent to 100 percent. No information was given on disparities in care, specifically.

We had some discussion in our subgroups about this. One issue that came up was it was unclear as to whether, if aspirin is contraindicated in a patient but they are on Plavix, does that mean the measure would have been met? Really, the only exclusion criteria speaks to if aspirin is contraindicated. So that was one issue that was raised within our subgroup.

In terms of scientific

acceptability, it is clearly a useful measure for consumers and patients, and the scientific evidence is strong.

In terms of usability, the measure provides useful information, but one issue that was identified in our subgroup was that it was noted it is a measure of one of 11 component measures of a CABG composite score. So we wondered if there was clarification about how the measure is treated within the composite score. For instance, is it weighted equally with all measures?

In terms of feasibility, the measure is easy to implement and track. So, really, that was all that we had with respect to discussion of the measure.

CHAIR MORRIS: Thank you. Does anybody have anything to add to that? Issues, comments, questions? Anybody from the work group? Okay. Would developer like a chance to respond to that?

DR. PRAGER: David, are you on the

phone about the composite? Dr. Shahian or Jane?

DR. HAN: Hi, this is Jane Han. I was having difficulty getting in.

DR. PRAGER: Okay. So David may be having the same.

DR. HAN: I actually told him to join us at 9:25, since that was the time on the agenda, but are we running ahead of schedule?

DR. PRAGER: Well, we didn't go through the lengthy review.

DR. HAN: I know he will be joining us in about 15 minutes. Sorry about that.

DR. PRAGER: There are a couple of questions the STS needs to address, or at least two questions. One is -- and I think we need David for this -- how this is weighted in the composite metric, which was one question that came out of the study group. I am not sure that, actually, I can answer that for

you.

The second issue about, if not aspirin, if allergic to aspirin, was it a question about Plavix? I'm not sure.

DR. KLEINPELL: Right. if the patient is on Plavix, does that consider that the measure is met, because the only exclusion criteria speaks to contraindications for aspirin?

DR. PRAGER: Right. My understanding is yes for that.

DR. WILHOIT: I think the question there perhaps is whether you count the numerator event first or whether you count the exclusion first, because aspirin sensitivity is listed as an exclusion, but taking an alternative drug is also listed as a numerator event.

So I think the question probably is what order you count things in, whether you take the exclusion first or the numerator event first.

DR. PRAGER: And that is a good question, and I am not sure I know how we sample that via the database, which we take first, and I understand your question.

CHAIR MORRIS: Jane, can you speak to that? Do you guys perform your exclusions first before you gather the numerator and denominator?

DR. HAN: I, unfortunately, am not the one who does the analyses. So I would have to check with our data warehouse, unless Dr. Shahian knows the answer to that.

DR. PRAGER: We will have to go to DCRI to find that out, unless David knows.

DR. HAN: Right.

CHAIR MORRIS: Okay. Is there anything else that anybody wants to bring up for this measure? All right.

Then just also to note -- Jane, I am not sure if you heard this, but another issue that arose was the question of whether disparities have been measured in the

application of this care. I just want to reinforce that. I know that it is not present in the documents from STS, and I am suspecting that, like yesterday, it could be done. It just hasn't been done.

DR. PRAGER: Correct.

CHAIR MORRIS: So I want to underscore that. If there is nothing else, let's go ahead and move on to the vote.

So the first vote: Does the measure meet NQF criteria for importance to measure and report? Twenty-one out of 21 say yes.

The second vote: Does the measure meet NQF criteria for scientific acceptability of measure properties? Let me ask you all to press your vote one more time, and press Send. Eighteen say completely; 3 say partially.

The third vote: Does the measure meet NQF criteria for usability? Twenty-one out of 21 say completely.

Then the next: Does the measure

meet NQF criteria for feasibility? Twenty say completely; one says partially.

Then lastly: Does the measure meet all of the NQF criteria for endorsement, and the issues that arose were the fact that there is indeed a gap -- so that is on the positive side. There are several other positives, and then sort of open questions are what effect does this particular measure have on disparities; secondly, if aspirin is contraindicated, is Plavix an acceptable alternative; and thirdly, how is this measure treated in the composite score with regard to weighting.

Then, let's see now, the fourth issue was -- What was the fourth issue? It was when are the exclusions applied? So it was just a question, really, when are exclusions applied, and a pretty simple question. I think, generally, they are probably applied before capturing the entire numerator and denominator.

Anybody want to bring anything else up or anybody want to discuss these issues further before we vote?

Okay. Does the measure meet all of the NQF criteria for endorsement? Twenty-one out of 21 say yes.

The next measure is Dr. Collins, 0118, the anti-lipid treatment discharge. It is being introduced by Dr. Collins.

DR. COLLINS: Sure. Good morning. I have both 0118 as well as 1479, which are very similar measures. I don't know if you would like me to present both or just one at a time. I think we have some harmonization, potentially, discussions here.

CHAIR MORRIS: What I would like to do is to have you go ahead and present the first one. We will vote on it, and then present the second one. Maybe we should talk at that point about competing harmonization.

DR. COLLINS: That sounds good. So 0118 is an existing maintenance measure

with the steward as the STS, very similar to the previous measure we just discussed, and it looks at lipid lowering agents following CABG therapy.

The simple numerators are patients who received lipid-lowering therapy at discharge, with the denominator patients on CABG. Patients are excluded if anti-lipid therapy is contraindicated or if there was an in-hospital mortality.

No comments, I believe, in the proposal on disparities of care. This is an existing measure. The compliance is, I believe, around 98 percent in what was reported, which is very high. So our work group did have some questions on whether this measure was tapped out, being at 98 percent.

The importance of this measure, I don't think, will require too much discussion. It definitely still remains a very important measure as far as outcomes data associated with lipid therapy and, really, the work group

did not have major comments on science, acceptability, usability or feasibility.

CHAIR MORRIS: Any other comments anyone one has about this measure?

DR. DUTTON: Sorry. It took a moment for the coffee to start working. But both this one and the last one: Has the STS -- since these are returning measures, has the STS looked at why patients don't get them, when they don't. In other words, have they analyzed the failures:

The question would be, are they preventable or not preventable, because if most of the failures are not preventable like patients on tube feedings going to a nursing home or absolute allergic contraindications or something like that, then there is no point in keeping the measure. But if the gap is preventable stuff like, oh, we forgot or they couldn't fill their prescription because they are poor or whatever, then, obviously, we should keep it. Does STS know?

DR. PRAGER: At the national level -- in other words, via DCRI and then the national population of patients -- the STS does not know. At the regional level where this is looked at and most of the quality initiatives occur, what has been seen is that there has been increasing utilization of it, either via order sets that demand it or demand the reason that you do not use it; and while 98 percent looks great, everywhere is not 98 percent.

So that is what we have seen. Is there a method to see why it is not, is it definitely contraindicated? Have we drilled down? The answer to that is no.

I am happy to anticipate your other question about, if it is at 98 percent, should we keep going? Was that the next one? Yes. We have actually talked about this, and I would understand -- I understand the question totally, and we asked ourselves the same question.

What we have seen and what we hope to really accomplish by keeping a measure such as this is to allow other people to come up with a system, so that this becomes part of the mindset of a postoperative medication, one, and it isn't there everywhere because there are regional differences, as we talked about yesterday, in things such as the process measure of alima, and we are concerned about slippage. So we would like to keep this.

DR. SHAHIAN: This is Dave. I would say that the vast majority, probably approaching 100 percent of our CABG patients, would fall into one of the categories for which that therapy is recommended by ACT and AHA, based on a fairly large body of evidence regarding secondary prevention.

Of course, there is now a lot of evidence in cardiac surgery that it is valuable preoperatively as well, which may end up being something we will bring back to you in the future, but I think we would very much

like to continue this measure.

CHAIR MORRIS: Any other issues that anybody wants to bring up with regard to this measure? I think that that is a pretty insightful comment, and I guess that, as a group, we would like to really encourage the STS to think about some of these measures. They are really excellent quality measures, but maybe topped out in the near future or just beyond the near future.

So understanding why particular treatments are not received would probably be very useful to know if those cases in which treatment is not received were actually preventable or should be changed.

If there is nothing else to say, let's go ahead and move on to the vote.

Does the measure meet NQF criteria for importance to measure and report? Twenty-one out of 21 says yes.

Next vote: Does the measure meet NQF criteria for scientific acceptability of

measure properties? Twenty say completely; one says partially.

Next: Does the measure meet NQF criteria for usability? Twenty say completely; one says minimally.

Does the measure meet NQF criteria for feasibility? Twenty-one out of 21 say completely.

Then lastly: Does the measure meet all of the NQF criteria for endorsement?

Is there, before we start the vote, anything else that anybody wants to bring up? So to briefly recap, there is a question of whether this is tapped out.

We know that there is still some regional variation based on what our representatives from STS have said. They strongly desire to increase the utilization, as has happened so far probably with standardized order sets or other things that make it very simple to order these meds.

Then we will be addressing in a

few minutes whether this is competing with the next measure.

So with that, does the measure meet all of the NQF criteria for endorsement? Let's go ahead and vote. Twenty-one out of 21 say yes.

The next measure, 1479.

DR. COLLINS: Sure. The next measure is, like we have mentioned, very similar to the previous measure, also looking at patients 18 years and older who have had lipid-lowering therapy following CABG, and the steward is a company named Ingenix, which I believe is on the phone for comment as well.

This measure uses pharmacy claim database where they look at lipid-lowering therapy either 90 days prior to CABG, seven days following DC after CABG, or a procedural code at discharge.

So I think that is the major difference. It is really looking at the pharmacy claim data from what I believe is

either a 15 million or a 65 million member database.

Exclusions are pretty much similar to the previous measure: Mortality; if there was a readmit within seven days to the hospital, or if patients drop pharmacy coverage or Ingenix coverage, I believe, prior to when the script was filled.

The work group thought that it, like the other one, was an important measure. Some of the comments came as far as, if a patient did not fill the script after discharge, would the hospitals then become accountable for that, and some of the inaccuracies maybe with using pharmacy claims versus self-reported measures, as with the STS.

I don't believe there were comments on disparities of care, and I was a little unclear as far as cost outside of patients who are under the Ingenix umbrella.

I will also point out that one of

the issues I think we need to discuss in this is the percent of patients who have CABG. I am sure -- I don't know the numbers, but I am sure it is 40, 50, 60 percent are greater than 65 years of age, which I question whether this measure would capture those patients.

CHAIR MORRIS: Thank you. Are there any other issues or comments anybody wants to make about this?

DR. MORTON: I had a question. Is the only way to get the data through Ingenix?

DR. BURSTIN: The measure specifications are freely available. Anybody could run it using claims data.

DR. WILHOIT: Having additional measures that can be run with an administrative dataset can be a real advantage. While this is similar to the previous measure, the difference, I think, or a major difference is that for a health plan or for a large provider group that gets feedback on their pharmacy claims or whatever,

there is the availability to run the data; whereas, a lot of the STS data is not necessarily available to outside entities.

So I think in many respects it is a very different measure, even though it is looking at the same thing, because of the different data source or the different availability of information. However, that being said -- and I think the second thing that is different is the difference between prescribing a drug and filling a prescription.

The STS measure that we just looked at had a mean of in the high nineties or mid-nineties. This one, the rate was 32 percent, 32.8 percent. Well, either one is wrong or the other is wrong or we have got a huge issue.

If 95 percent of people are really being prescribed drugs and only 33 percent of people are filling the drugs, then we are fooling ourselves to look at the STS measure.

On the other hand, this 35 percent

really seems unrealistically low, and that makes me wonder if there are problems with the measure. So I think that adding measures that can be run using administrative data is important, but it seems like there must be a disconnect here.

Coming back also to the issue of whether the data could be run for patients age 65 and older, there are many retirees who are still covered under their employers' or former employers' health benefit plan, and a lot of people who are continuing to work after that age, and it is usually clearly identified in the administrative dataset whether somebody has pharmacy benefits and, if they don't have pharmacy benefits, I believe they are excluded from the measure.

So I think that particular issue isn't of particular concern.

DR. SAIGAL: Can I comment? I agree. I see two points. One, Carol's point about the low rate of filling -- I do a lot of

work with claims. There is a lot of noise when I look at those claims data.

I was wondering if there is any validation studies done on this measure looking at whether patients actually who didn't get a pharmacy fill in their claims database got a prescription, if we can do any clinical correlation with that on a small scale, and also how do they deal with exclusions that are clinical in nature like on reactions to Lipitor, something like that, in a claims database.

DR. CIMA: Also to follow up on Carol's point, when we have looked at this in our institution about what people right after surgery, not filling their prescriptions right away, oftentimes there's confounders into that.

So like I just got out of the hospital, and my husband is also on Lipitor, same prescription; I started taking his. I am not going to fill it until I feel better.

There's all sorts of weird issues.

The other thing, the fundamental issue and my main concern with this is attribution. Who is going to be responsible for this? So who is going to get the -- When you do public reporting on this, what is it going to say? Is it going to say hospital A only performed at a certain level on this, when they had no control on whether or not that patient fills that prescription?

I have real serious concerns about the quality of the data as far as the amount of lives covered, and to Carol's point, why was there only 30-some-odd percent of patients saying they had this? Is that really the gap?

You know, even if we take the STS as a rosy picture, this would be saying that we are doing a terrible job. So my main concern is attribution. How are you going to attribute who is responsible for owning this and saying we can make it better? Is there really a quality improvement initiative that

a hospital can do if patients aren't filling the prescriptions?

There are all these other ways around handling administrative pharmacy data, and there is a lot of noise in it. I think we have a cleaner measure with the STS one. This one doesn't really add a lot of value as far as quality improvement, and it is going to make public reporting somewhat of a nightmare for institutions to try and handle.

MS. STEED: It is actually not clear how they are going to use it for public reporting.

DR. HALPERN: I also do wonder about the age issue, because they do say this database represents a predominantly commercial population less than 65 years old. So what percentage of their patients are 65 years and older that they are actually analyzing, since again, like somebody else pointed out, people who get CABGs are generally over 65?

DR. MORTON: I had a technical

question about would this measure only work if you have a pharmacy benefit? If that is the case, I don't know how often people don't have a pharmacy benefit for this particular surgery.

DR. WILHOIT: In a commercial health plan setting, it depends on the health plan. However, depending -- For us, depending on the product, it ranges from about 40 percent of members with a pharmacy benefit to about 85 percent, depending on the particular kind of product.

Not having a pharmacy benefit is an exclusion from the measure. So that, you know, it is accounted for. The other thing -- and I can't speak for Ingenix, and if they are on the phone, they may be able to respond, but in terms of the database and whether the people were under 65, I think that was the database in which they did the analysis, but the measure, I think, would be intended for use in other databases as well.

CHAIR MORRIS: Dr. Dutton, did you have something to add?

DR. DUTTON: Well, I was going to say, if you have to have a pharmacy benefit to be included in the measure, isn't there an inherent socioeconomic bias in the data already that is going to make it very hard to use this data for looking at disparities.

DR. WILHOIT: The other side of that is, if one is, for example, a health plan or one is an integrated delivery system and you are trying to look at your own data, you know, the administrative data is what you have, and that is what you can work with.

The other advantage of using pharmacy claims -- I know from our experience, we have pharmacy data pretty complete within a month; whereas, claims data for other kinds of services is three, four, five, six, eight months, depending on what you are looking at, and STS is a whole lot longer than that.

So one of the real up sides in

terms of things like identifying gaps in care,improving gaps in care, is that this can be assessed on a very timely basis. So that is a real positive as well.

DR. MORTON: I guess my only concern is, if pharmacy benefits are an exclusion, you are going to leave out anywhere between 15 to 60 percent of people that are undergoing the procedure potentially.

CHAIR MORRIS: So a lot of different issues arise with that. Does anybody have anything else before we give Ingenix an opportunity to say a few words, and also I would like to just let the folks on the line know that they certainly can have a little bit of extra time, since they were unable to introduce their measure, because our phone lines were not open.

Any other issues before Ingenix responds? One more?

MS. ZAMBRICKI: I would just like to speak for the fact that this is a big

concern when I read it, this difference, and we are not sure exactly what it means. But I would hope that, whatever the decision is, that somehow this continues to be measured and some attention be paid to it.

I think the attribution issue is an important one, but as a global public health issue, I think this is really an important question.

CHAIR MORRIS: Okay. So just to recap, a lot of different things came up. One question is what percent of patients over the age of 65 years old would actually be captured using this system? How will this be used for public reporting is unknown, as I understand it. There are issues with attribution or accountability at the hospital level, who is accountable, particularly if patients elect not to get their prescriptions filled.

There is no information about disparities, and it seems unlikely that using this measure we would be able to obtain a lot

of information about disparities, because those without a pharmacy benefit would not be captured.

There is a question about the cost burden to hospitals, which was unclear, and a couple of related questions. Using administrative claims, how does Ingenix deal with the noise in this? How do they address exclusions using this measure, and why is the measure uptake -- why does it appear so different from measure uptake in the previous measure?

One of the questions with regard to that was the fact that, if patients already have a prescription for statins or lipid-lowering medication at home, would they not be captured by this measure? So this specifies taking a lipid-lowering medication at admission or within seven days of discharge. It is possible that they may have statins at home that, for some reason, are not captured, and that may be why there is such a low rate.

Dr. Halpern?

DR. HALPERN: I have also noticed that one of their -- Included in the measure is 90 days preop. So I am wondering if that is the difference in the percentage, because maybe people don't have it prior to coming to the hospital, which is an important issue also; because as somebody mentioned, if it is prior to surgery, both cardiac and vascular seem to help with overall morbidity and mortality from the surgeries.

CHAIR MORRIS: Okay. Would our Ingenix representative like to respond to that?

DR. RIEHLE: This is Jessica from Ingenix. Can you hear me?

CHAIR MORRIS: Yes.

DR. RIEHLE: Do you want me specifically to respond to using Lipitor or lipid-lowering medication at home prior or do you want me just to go through the list that you read off?

CHAIR MORRIS: I would like for you to go through the list.

DR. RIEHLE: Okay. So there was a concern raised about the percentage of patients 65 and older. Our database specifically that we use to test the measure does not have very many people who are over 65. However, there is nothing inherent to the measure itself that ruled out patients who are older than 65, and a lot of our customers have data for patients over 65.

So the measure still applies to that population. Unfortunately, with our database we weren't able to test it in that population, but there is nothing that would exclude older patients.

In terms of public reporting, this measure is being used for public reporting, mostly at the physician level for provider measurement.

In terms of attribution, you know, we don't have specifications as to how the

measure is attributed. The people that use our measure -- that is something that they define.

We also share the concern about patients who don't fill their medications, which is why we included in the numerator a G code, which is a code that a physician can use to say that they prescribed the medication, and it is not at all dependent on whether or not the patient fills the medication.

In terms of exclusions for people who might have an intolerance to the medication, unfortunately, that is really hard to do with claims data. There really isn't a great way to code the fact that somebody may have a history of intolerance to the medication.

For the Lipitor prescription at home, again the numerator does include patients who filled a lipid-lowering medication during the 90 days prior to the CABG admission. So people who may have the

medication at home -- you know, if they are taking their wife's prescription, unfortunately, there is no good way to capture that using claims.

CHAIR MORRIS: Can you clarify your thoughts regarding why that measure uptake appears so different from the STS measure uptake?

DR. RIEHLE: In terms of the compliance, the 32 percent versus the 90-whatever percent?

CHAIR MORRIS: Yes.

DR. RIEHLE: You know, I am not sure. I suspect that -- I mean, I would be very, very surprised if the compliance was as high as 90-something percent. We would like to go and actually do a comparison eventually of our data versus electronic charts or even paper charts. That is something that we would like to do soon, but we have never done that. I am not sure why you see such a discrepancy, to be honest.

DR. SAIGAL: Can I ask a question? Could you not also use a G code to eliminate people that are intolerant to these medications, if you are using the physician reported one?

DR. RIEHLE: I am not sure if there is a G code. There might be a CPT-2 code. There probably is some sort of a code, and we could definitely look into that.

DR. WILHOIT: There is a code listed in the denominator exclusions on page 7 under QA-10. There is a G code, 8586, which is anti-lipid treatment contraindicated/not indicated.

DR. RIEHLE: Oh, okay. So it is there.

CHAIR MORRIS: It doesn't mean that the providers will know that that is there.

DR. RIEHLE: Yes.

CHAIR MORRIS: Are there any other issues anybody wants to bring up before we

vote?

DR. STAFFORD: Yes, I have a question. Did I understand you correction that public reporting would be at the physician level?

DR. RIEHLE: You know, I mean, it could be used in a variety of different ways, but that is primarily how it is being used now in terms of public reporting, would be at the physician level.

DR. STAFFORD: So I would -- This gets back to attribution and attribution bias. There is a huge problem with that, particularly in academic centers where a prescription for a medication might get written by the resident, and so if you are looking at attendings, then it wouldn't show up as having been written by the attending. I think that could be a huge problem.

DR. HALPERN: Not only that, if the doctors are responsible for the ones putting in the codes, if it is a resident

putting in the -- A resident won't be putting in those codes.

DR. DUTTON: I will pile on that also. We are trying to encourage team practice and accountable care and bundling of episodes and so on. Attributing data like this to individual physicians is just horrible for that, because you don't necessarily know who the responsible doctor is, and in an appropriate system it might be an internist who is managing that patient's medications through a surgical episode.

CHAIR MORRIS: I believe that the accountability problem here also resides in the other measure. If this is to be published at the hospital level or publicly reported at the hospital level, that would match, presumably, the other measure from STS.

If it is to be reported at the physician level or whoever it is that is measuring it decides to report it at the physician level, then obviously that group is

really concerned about that.

DR. CIMA: But even at the hospital level, there are things out of your control -- you know, what plan they have, whether their plan is covered. It poses a risk to the hospital, even it is on the dismissal summary.

We should ask people to do what they can do, not to ask them to be responsible for the world. I think this attribution issue is a major issue, and institutions have to be sensitive to it, and we have to be sensitive to that also.

DR. SEARS: Yes, I guess we are all piling on about the attribution issue. I think one thing, we pass a measure like this, hospitals will rethink what they do.

They may have to actually give the prescriptions to the patients so that, when they go home, they know that the prescription has been filled, and then they have satisfied their obligation to the measure, and in a DRG

world where 65-70 percent of these patients are probably Medicare, they are not going to be able to get any collection for the drug that they are going to have to give out.

DR. SAIGAL: I just have one last comment. I do think that the issue of the difference in the rates between this measure and the STS measure needs to be looked at before this measure gets put through, because I think a small validation approach to what they are doing would be really helpful and help me believe that this is going to be useful in public reporting.

DR. WILHOIT: And I totally agree with that. For me, that is the biggest issue here. I think adding some administrative measures is really positive. I think there's a lot of things that are positive about this, but at the moment, for me it lacks face validity. Thirty-five percent just seems -- just doesn't fit the sniff test.

That makes me wonder if there is

some of the logic that isn't quite correct. So for me, it needs some further evaluation and, if the rate really is this low and we are kidding ourselves with the 95 percent, then that is really worth knowing, and that is very important; because if we want good outcomes, we need to make sure care is actually delivered. But I think it needs testing to try to understand that and make sure it is not a logic error.

CHAIR MORRIS: Helen?

DR. BURSTIN: Just two comments, one of which is: There is actually very clear and known literature of the low rate of compliance with statins post-discharge. I mean very low rates. Thirty percent is actually what people tend to say for people actually on statins beyond six months.

So it is actually hard to know which is actually correct. Ninety-eight percent is probably true in terms of saying, yes, please be on a statin at discharge. It

is very different to say a patient actually went, took the prescription, and filled it.

So they really are measuring very different concepts, and we need to better understand it. My preference personally is to go to the one where we actually know the patient has got the drug in their hand or, even better, skip that entirely and just look at LDLs, which is really the end test here of are you on a statin? Are you taking it, and is your LDL in control? Neither of these kind of really get at what I think is truly the end game here.

Just lastly, just because this comes up on every single Steering Committee, this issue of accountability and attribution is just a really difficult one. The reality is we need to pick the measures that we think are best to serve the needs of the public, to get to the right assessment of quality.

We are really trying to move toward models of shared accountability. It is

not just the clinician. It is not just the hospital. It is not just the pharmacy who fills it at the end of the day. But the only way to do that is to pick whatever the best measure is, and the attribution issues, I think, are just going to -- will always make us take a step back from potentially measures that would really drive improvement.

We would potentially not have done readmissions. We would not have done -- There is a whole series of things we have been able to make improvements, because we kind of took the step toward the tougher measures. So I am off my soapbox. Thank you.

CHAIR MORRIS: Okay. I think it is time for us to move to a vote.

DR. SHAHIAN: Excuse me. This is Dave Shahian. Is it permissible for me to make a comment as somebody involved with the other measure?

CHAIR MORRIS: Sure. Go ahead.

DR. SHAHIAN: It strikes me that

this is really a completely different measure in many respects. Well, one of the most important that I see is that the measure would be satisfied, as I read it, if one were on a lipid-lowering medication at the time of CABG admission.

Now that, clearly, is out of the control completely -- I know we just talked about this a second ago, but it is totally out of the control of the surgeon, and the surgeon could -- Our measure is trying to determine whether surgeons and their team, including cardiologists, are giving a statin prescription or a lipid-lowering prescription at the time of discharge.

This measure would be satisfied, I think, if a patient simply came into the hospital on a lipid-lowering medication. Am I correct about that?

DR. RIEHLE: Yes.

DR. SHAHIAN: So that strikes me as a completely different measure. I am not

saying whether I favor it or not, but I think it is a much, much different measure in many respects.

CHAIR MORRIS: Thank you. Are there any other comments? Dr. Collins, can you speak to the discussion among the work group regarding whether you felt that this was a competing measure or whether it was substantially different from the previous measure?

DR. COLLINS: I believe the work group thought that they were competing measures, and the question, like I mentioned before, of harmonization or we were a little unclear, if we had to pick a winner, of what our course was there. But we thought they were competing and not completely separate.

MS. MURPHY: And the requirement that you had before you is to evaluate this measure with its specifications, and the discussion about harmonization/competing can follow later.

CHAIR MORRIS: Thanks for clarifying that. Let's move on to the vote.

Does the measure meet NQF criteria for importance to measure and report? This speaks particularly to impact, performance gap, and evidence. I will ask everybody to press their buttons one more time, and hit Send. Twelve say yes; nine say no. So we will go ahead and proceed.

Does the measure meet NQF criteria for scientific acceptability of measure properties? One says completely. Seven say partially. Twelve say minimally, and one says not at all.

Does the measure meet NQF criteria for usability? Three say completely, six partially, nine minimally, and three say not at all.

Does the measure meet NQF criteria for feasibility? I will ask everybody to hit their button once more, and hit Send again. One last time, and if you notice that you are

consistently potentially the last voter, please see me at the break. We will change your battery or something like that, the battery in your voting item, not your personal battery. Five say completely. Eight say partially, seven minimally, and one says not at all.

Then the last vote is: Does the measure meet all the NQF criteria for endorsement, and I would like to recap some of the major issues that were brought up.

There was a lot of question about the validity of this data compared to the -- or of this measure uptake compared to the measure uptake for the STS measure, and a lot of concern about the big gap there with questions about which one could potentially be more accurate or whether they are really measuring different things.

We heard from Ingenix that at some point they may have a plan to validate their claims method by comparing to chart derived

data. That sounded, to me, a little bit fuzzy, as had been discussed, but that a plan doesn't exist at this time.

There were issues regarding capture of -- adequate capture of patients, particularly those who don't have pharmacy coverage, and whether or not we would be able to learn anything about disparities in care using this measure.

There were issues about attribution accountability at the hospital level, at the physician level, holding folks accountable or institutions accountable that really had no control over this outcome.

There were questions about how exclusions were dealt with. A lot of times Ingenix said that the exclusions couldn't actually be addressed using claims data, but it sounds as though there are some claims that indicate when patients are unsuitable for use of lipid-lowering medication, or it sounds like that claims are not known by many

providers potentially.

There is a question about how to deal with the noise that is inherent in claims data, and to my mind, that wasn't truly addressed, but it may not be possible to completely deal with the noise in claims data. On the plus side, claims data is pretty easy and cheap to acquire.

There were issues about cost burden to hospital. That pretty much summarizes it for me. Does anybody else want to bring anything up with regard to this measure?

DR. SIPERSTEIN: I just want to comment. I think the goals of this measure are very laudable in that it looks at the next step after we write our prescriptions. It starts to look at the whole issue of patient compliance, and it is, obviously, part of physicians' responsibility to educate their patients to the importance of filling their prescription and taking their medication.

I am just not convinced that the measure as written really serves that goal. We get a hint of that when we do our medication reconciliation when a patient comes back a week later. It is not a perfect system, but I would encourage the authors to continue to work on some similar measure, because I think this is an important thing to look at.

CHAIR MORRIS: Thank you for making that comment. That also speaks to Dr. Burstin's comment about what is the real outcome that we are going for here. Are we going for the outcome of just prescribing the medication or recording that one has been prescribed or are we going for the outcome of patients actually taking the medication or the end game, which is better health or lower LDL?

So I think those are important to keep in mind with all of the measures.

DR. STAFFORD: Along with that point, Dr. Dutton mentioned the potential for socioeconomic bias, and those are exactly the

people that we probably could help the most, and they are being excluded from this database. So if they have trouble getting them filled, there is no way we are ever going to capture that with this database.

CHAIR MORRIS: Thanks. So let's move on to the last vote. Does the measure meet all of the NQF criteria for endorsement?

We have one yes, 19 no, and one abstaining.

We will move on to the next measure, which is Measure 0130, deep sternal wound infection rate by Ms. Steed -- introduced by Ms. Steed.

MS. STEED: Yes. This measure is an established measure already, and it is the percentage of patients age 18 and older undergoing isolated CABG who within 30 days postoperatively develop deep sternal wound infection involving muscle, bone and/or mediastinal, requiring operative intervention.

It has a pretty clear numerator

and denominator statement, and in the discussion with the group there was really no significant conversation about this measure via the importance of scientific, usability or feasibility except for one comment.

That is, at the present time I understand that are two organizations, the CDC and the American Academy of Surgeons, who have proposed surgical site infection definitions to NQF, and I understand that they are in the harmonization phase. They have not harmonized those definitions, but when those get approved, then this particular measure will have an issue related to harmonization.

I don't know what the American College of Surgeons' definition is, but I know what CDC's definition is, and this particular definition differs in that it looks at the infection developing within 30 days of hospitalization or the surgery, to where CDC's goes up to 12 months postoperatively, and that is the biggest difference.

DR. BURSTIN: The CDC measure goes longer out only if there is an implant.

MS. STEED: If there is an implant. You are right, if there is an implant, and they consider sternal wires implants. I hate to say it, but that is the truth.

DR. BURSTIN: It is actually 180 days, but still that is a good point. I don't know the answer of whether that --

MS. STEED: Yes, sternal wires are considered implants by CDC, which is one of the controversies between, I am sure, the American College of Surgeons and the CDC.

DR. SIPERSTEIN: My understanding of reading that infection measure is that sternal wires would not count. They are talking about joints, valves, but not wires, which are variants of sutures.

MS. STEED: Being someone that has to conduct the surveillance for CDC and being involved in public reporting, and I am in the

state of South Carolina, when we get validated they do consider -- CDC considers sternal wires as an implantable, which is one of the biggest controversies that surgeons have with their definition.

So just know that that is the case, and I am sure that is part of the harmonization that is going on between the American College of Surgeons and the CDC, but I do not know where that stands. I am not involved in it, but I felt it important to bring it up.

DR. BURSTIN: It is actually a good point. I believe part of the harmonization effort to date has been to at least take staples out of the definition, but I don't know about wires. Staples was actually considered.

MS. STEED: I know it was -- guide wires -- Put staples in there, and you have to follow a guide wire for 12 months.

CHAIR MORRIS: Any other issues

anybody wants to bring up? Okay. Would the STS like to respond?

DR. PRAGER: Yes. We recognize the differences with the aspects of the CDC definition. I am not sure, frankly, we were aware that wires are implants, but they apparently are.

At this point, as we have said over the last day and a half, we do not have measures that go out to a year at this point in time, which is what the CDC does with implants. Ideally, you would love to know these pieces of information, but the practical side of this at this point is that it is not being done.

DR. SHAHIAN: This is Dave. I would just add that we spent a lot of time on this particular one this year, and the specification upgrade. There were a few minor differences between our measure and the CDC definition, and we did, in fact, make those changes in order to make it completely

consistent, except for the 30-day versus one year, which is simply impractical for us to implement at this point. But in all other respects, the measure is now consistent with the CDC definition.

In fact, although there a are very, very small number of smoldering sternal infections that occur late, I would say that the vast, vast majority occur within that 30-day window.

CHAIR MORRIS: Thank you. I am actually curious about that small number of smoldering infections. In colorectal surgery, we know that with a colonopy anastomosis, about 12 percent of them occur -- become apparent after 30 days. We know this based on pretty good registry data.

So measuring anything up to a 30-day window always leads you to wonder what is happening after 30 days. Do we have any hard numbers at all regarding what happens with sternal wound infections?

DR. PRAGER: I am not sure David or I -- I don't want to speak for him -- have hard numbers. There have been a couple of anecdotal case reports in the literature of patients occurring -- having mediastinal infections late, and I don't mean day 31. I mean three months, five months, six months, and frankly, we have all seen it.

Where that is, though, percentagewise, I frankly do not know. David, do you have anything to add?

DR. SHAHIAN: As you say, there are a very few reports about this, and I don't have them at the tip of my fingers, but the number is really quite small.

DR. DUTTON: Just a science question for the cardiac surgeons. Are these ever managed with percutaneous drainage or nonoperative treatment? I know open exploration is the recommended approach, but do you think you miss some in the numerator, because the patient is very sick or for some

other reason that are managed nonoperatively?

DR. PRAGER: That is a good question. I would expand it a little bit to say that, if a wound vac is now placed, which is now being done not infrequently, that is considered an operation, and we are capturing that. At least, in the new specifications, we will.

There are opportunities for percutaneous drainage via interventional radiology usually. Our experience with that has been that has occurred even after the exploration, less likely to take the place of an operation, but I wouldn't say that my statement is 100 percent.

DR. SHAHIAN: I would say that, unlike an intra-abdominal abscess that may occur after colon surgery, for example, which can be -- if there is no active leak, can be treated with drainage an antibiotics, I don't think I have ever seen a true sternal infection/mediastinitis effectively treated

without reopening the sternum and doing something.

Now you may reopen and put -- Some people have put drainage and irrigation tubes and used various agents to irrigate the mediastinum. People use vacs. people use flaps, but to treat it completely percutaneously -- never seen it in 30-plus years.

MS. STEED: Another comment I wanted to make is that CMS is going to be utilizing CDC surgical site data at some point for public reporting and reimbursement. In doing so, CDC's definition doesn't only include deep surgical site infections. It includes superficial, incisional and organ space. So, therefore, the surgical site infection rates that will be reported via CMS, via the CDC, will be higher than the rates reported by this particular metric.

DR. SHAHIAN: We also capture the superficial separately.

DR. WILHOIT: One thing I wondered in looking at the measure is how useful it is, other than as part of the composite. According to the materials, the rate is about a half a percent, which means that you would have to do about 200 cases to have one infection on average.

You know, from the data we saw yesterday, a lot of facilities or practices are not above that 200 mark. When you look at the distribution of results, they show that out of 640 groups that were assessed, there were 54 outliers, so a little less than 10 percent outliers. Of those 54 outliers, 53 were low.

When you look at the distribution, there were a lot of zeros or near-zeros, probably because of the adjustment methodology, and there was only one high outlier.

So is this even useful? You can identify the people who have a rate of zero

and come out low. Well, that is this year, just -- which, because of the small numbers, may be chance, but there is very few high outliers identified. So is this even useful?

DR. SHAHIAN: Well, there, historically, and even today, I'd say, is a five -- probably at least a fivefold, if not greater, variation in the prevalence across institutions. There are institutions that have reported anywhere from zero to .3 percent deep sternal infection rates over a period of many years. There are institutions that have reported rates of two to three percent.

So there is variability, and I think this is one of those measures where there are some very well described interventions that can reduce the incidence of sternal wound infections. So I think there is a real opportunity for improvement, and there is a link to process measures that have demonstrated efficacy.

CHAIR MORRIS: Dr. Prager, do you

have anything to add to that?

DR. PRAGER: No.

MS. STEED: I think that there is significant morbidity and mortality associated with this particular metric, which is the reason why I think it is important.

DR. PRAGER: Yes. This is a catastrophic complication, and if you put it in the world of cardiac surgery with certain groups doing many immunosuppressed patients, more people looking to do two internal mammary arteries, I think we need this.

CHAIR MORRIS: Okay. Any other comments? We will go ahead and vote.

Does the measure meet NQF criteria for importance to measure and report? Twenty-one out of 21 say yes.

Does the measure meet NQF criteria for scientific acceptability of measure properties? Twenty say completely; one says partially.

Does the measure meet NQF criteria

for usability? Nineteen say completely; two partially.

Does the measure meet NQF criteria for feasibility? Nineteen say completely; two say partially.

So just to recap our discussion, we talked about potentially competing measures. We talked about the difference in the CDC definition versus the -- or some CDC definitions versus the STS definition. We learned that STS has worked to harmonize as much as they can the definitions, although they are still slightly different.

We learned that, although sternal wound infections don't occur that often, they primarily occur before the 30-day window, and that they are devastating when they do occur.

Anybody want to add anything to that? Okay, does the measure meet all of the NQF criteria for endorsement? Twenty say yes; one says no. Great.

Dr. Shahian, are you still on the

line?

DR. SHAHIAN: I am.

CHAIR MORRIS: We have a couple of questions from previously with regard to measure 0116. Let's see now. Who was it that introduced that? Dr. Kleinpell, would you like to?

DR. KLEINPELL: Sure. Our group had two comments or two questions that we wanted to identify or have questions on with respect to 0116, which was anti-platelet medication at discharge.

We noted that this measure is part of a composite reporting measure within the CABG composite score, and we wanted to know how is that measure treated within the composite score? For instance, is it weighted equally with all measures?

DR. SHAHIAN: In the composite, there are four domains: Risk-adjusted mortality; risk-adjusted morbidity; use of the IMI; and adherence to guideline recommended

medications.

Within the medications domain, there are four preoperative beta blockade and discharge beta blockade anti-platelet agents and anti-lipid agents. That domain, just as the morbidity domain, is a -- It is an all or none.

So to get credit for that domain, you need to prescribe all those medications, or you fail. However, in terms of the weighting among the domains, they are not -- They have equal weight, although because of the rather tight distribution of mortality scores, mortality ends up being, by far, the most important component, just by virtue of the standardization process. But there was no attempt to assign greater weighting to one or the other domains. Does that answer your question?

DR. KLEINPELL: Yes, thank you. That was helpful. We just weren't sure of that.

DR. SHAHIAN: Sure.

DR. KLEINPELL: Then the other issue: One of our reviewers mentioned that the exclusion criteria only really speaks to the contraindications for aspirin. So if a patient is on Plavix, would the measure have been considered met?

DR. SHAHIAN: Yes.

DR. KLEINPELL: Okay. That is what we thought. So thank you.

DR. WILHOIT: And related to that, that is something that is not clear in the document, whether you take the numerator event first or the exclusion first. So that is a slight improvement that could be made in terms of the documentation.

DR. SHAHIAN: We will note that. Thank you.

CHAIR MORRIS: Okay. So at this point, we are ahead of schedule, which I hope will last, but who knows. So let's go ahead and take a break until 10:30, and I will see

you all back here at 10:30.

(Whereupon, the foregoing matter went off the record at 10:14 a.m. and went back on the record at 10:37 a.m.)

CHAIR MORRIS: We are going to go ahead and get started here. The next measures are going to be discussed by our representative contractor with CMS. These are all maintenance measures, and the first one is 0300, introduced by Steve Findlay, cardiac patients with controlled 6 am postoperative serum glucose.

MR. FINDLAY: So this is measure 0300 titled cardiac patients with controlled 6 am post-op glucose. This is a hospital process measure around the issue of lowering the risk of infection associated with hyperglycemia for both diabetes and non-diabetes patients.

The numerator is surgery patients with controlled 6 am glucose below 200 on post-op day one and two. The denominator is

cardiac patients with no evidence of prior infection. There are quite a few exclusions. I won't go through them. They are in your paperwork.

The measure steward is CMS, and the measure has been in use since 2001, and it is used interactico and has been since 2007. It is also used as an accreditation measure by the Joint Commission, and the measure is going to be retooled for EHRs in the next year or two.

For the last two years, the measure score on this has been 90 to 95 percent in 2009-2010. Disparities were not assessed.

We had a really lively discussion on this measure on the work group call. Several people took issue with the measure's importance, clinical importance, the usefulness and design. I am not a clinician. So I can't respond to those issues, but I would invite particularly Bob and, I think,

Ruth raised some issues around whether this measure -- whether the 6 am value is indeed the best assessment of this. So I would invite those comments.

There was also concern about the measure being vague and just generally poorly designed at this point.

DR. KLEINPELL: I think some of the things we highlighted in the call was that it is difficult. If you have an early surgical patient come back, you have more time to rectify elevated glucose levels versus a later surgical day patient.

I think we have seen clinically an increased use in insulin drips in patients just to try and get their glucose to be below 200 the following a.m. to meet this criteria, and with increased use of insulin -- IV insulin therapy, we have had some cases of hypoglycemia, and the literature clearly indicates that even one case of hypoglycemia can increase hospital patient mortality.

So I think there are some issues with trying to meet it at 6 a.m. I think, clinically, we see from experience at our setting and other settings and talking to other clinicians, it is not necessarily the first day.

It is the second day when they are off the insulin drip, you know, to really try and keep them euglycemic, but I know this measure has been used now for several years, and everyone tries to achieve it. But it is intensive in terms of labor, you know, to be able to do hourly ECU checks and to keep patients in range. So it is labor intensive as well.

DR. CIMA: I think, from our point of view, my point of view, that is one of the main problems, is just the structure of it. You know, with the skip, one, two and three, with the antibiotics, we say 24 hours from some point, but in institutions that are doing high volume cardiac surgery, there is a huge

difference between a patient that is first case in the day and one that comes out at seven, eight o'clock at night as far as that 6:00 a.m., and it is not the way it is designed.

It is not the way the abstraction is done. It doesn't necessarily mean to be 6:00 a.m. It could be the 3:00 a.m. one, and then the next one is at 9:00 a.m., and you take the 3:00 a.m., but if it were a person that just got out of the OR at 9:00 o'clock at night.

So are you comparing apples and oranges? So that is a real -- It is not the -- The goal is good, although there is now a lot of data that says this probably isn't the best measure. Intensive insulin therapy has only been really shown to be effective in critically ill patients, and even then that is up for debate.

So whether it is actually a measure that actually does anything is another

story for two time points over a 48-hour period. It should be maybe a consideration of an aggregate measure of insulin control, but certainly, the way it is written is very vague. It makes for a lot of heterogeneity in the data that you are comparing.

That was my main concern from the get-go from this when it was first introduced, is that it is just poorly designed to find what you want, because you are comparing a lot of times apples and oranges.

DR. MORTON: I want to add to those comments, because what we see a lot of times in practice is people rushing around just to get that 6:00 a.m. value, and some other care doesn't always get rendered. So it is the arbitrary part about the 6:00 a.m. that bothers a lot of people in terms of implementing logistically.

DR. DILLON: Is this -- For those of you who have to hit the target of 200, is this going to change in the immediate time

period? I know there has been some talk about loosening how tightly controlled they have to be in the postoperative period. So are we going for an arbitrarily too harsh a measure here?

DR. KLEINPELL: No. You do want it less than 200. In fact, less than 150 is really recommended in cardiac surgery patients. I think the issue we are looking at in the literature is: Is glycemic variability a better indicator than just one isolated 6:00 a.m. glucose level?

DR. DILLON: Right, but as you point out, the issue of the hypoglycemia and the risk in terms of the population management is of growing concern, at least in our institution.

DR. STAFFORD: So the hypoglycemia was actually seen with what is classically described as intensive insulin therapy that came out of the Vanderburg study with less than 110, which is why I think the nice thing

about this measure is that they did choose 200 as opposed to 110, so that you don't get into as much trouble with the hypoglycemia.

I think you will find most institutions have gotten away from that 110, even for all of their other patients, because we have learned that that was a problem.

DR. DILLON: But the problem with that is that the 200 number is an arbitrary number, and it has not been shown to be effective. What is the difference between 210 and 190? There is no science that says that is a difference.

DR. CIMA: If you are chronically above that number in the hospitalized surgical patient, that is a problem. And at 6:00 a.m. the morning after a CABG, you know, you don't adjust for patients who are still on inotropes, which increase blood sugar levels, no matter what you give them.

So it was a poorly designed measure from the get-go, and it has not

improved, and I am really wondering if there is any evidence to support that it has made a significant difference.

MS. STEED: To comment, in my organization, even though I agree with the comments about using the 6 am glucose, I think you just take the blood sugar closest to that time frame. In our organization we started with the SIP measure and that initiative back in the early days in the early 2000's.

We saw 50 percent reduction in our sternal surgical site infection rate by controlling glucose, and can we prove it was that? Maybe not completely, no, but the perspective of the cardiovascular team was that the glucose control had an impact on our infection rates.

DR. CIMA: Did you also standardize the antibiotic dosing and the one hour before and everything? There is no other published literature that supports what you just said.

DR. HALPERN: And also, was it overall glucose support? Their main point is it is two arbitrary readings as opposed to total glucose control, and it is total glucose control that really makes the difference, not just two arbitrary readings.

MS. STEED: I agree with that.

DR. KLEINPELL: It is clearly a significant clinical issue. You don't want to have hyperglycemia in your critically ill patients, and I think this is less than 200. Really, you do want it less than 150, and many ICUs, regardless of if they are cardiac surgery patients or not, have developed insulin -- intensive insulin therapy with certain ranges.

We used to have 80 to 110. We moved it to 80 to 120, and now for our cardiac surgery we are up to about 150. So, certainly, it is clearly of clinical significance, but I think with this measure there are some issues in terms of usability

and, really, what is the impact.

DR. HALPERN: I think we are saying the same thing. I am basically saying it is overall glucose control rather than two arbitrary points which may or may not actually capture -- because if they are 500 the rest of the day, you are not really fixing them.

DR. STAFFORD: Yes. A better measure might be X percentage of blood glucose values below whatever. You are not going to find data for that, but it might be a more useful way to measure, because that would get at how well controlled you are for that entire period of time.

The other thing that I find interesting about all of this is that there is nothing being said about what blood glucose they come in with, and we all know that well controlled diabetes with hemoglobin A1c levels that are in the normal range before somebody gets operated on has an effect on outcome as well, and many of these cases are elective

cases.

So I would encourage people to start thinking about actually moving this kind of a measure back even further in the preoperative care of patients.

DR. DUTTON: I will comment on that as well. From the anesthesia perspective, the glucose control should start when we first see the patient and should be continuous through the operation, recover, and to the intensive care unit. So the time point is, I think, an arbitrary or pragmatic decision to make it easy to measure.

It is looking for your car keys where the light is good, because we can get that data easily, but there is no question that control should be continuous.

DR. CIMA: And, clearly, the evidence supports exactly what you said. Patients who are known diabetics who come in with A1c in the acceptable range -- their postoperative morbidity is less.

So a better measure, if you really wanted to make population improvement, would be to say people with known diabetes, you don't operate on them until their A1c is in a certain level unless it is an emergency. But that is not what we are faced with.

We are faced with a very poorly designed measure that was an attempt to get people to do insulin therapy, but it doesn't support -- The science doesn't support this value. It should be lower, which is not necessarily practical or safe, necessarily, in some cases; and it is very arbitrary in how it is designed, and doesn't take into account the heterogeneity of the population in which it is being applied.

If everyone did one CABG a day, and that patient got out and got to the ICU at Noon, then I would say it is reasonable to go to 6:00 a.m. as your first marker, but other than that, it doesn't seem to pass sort of -- It is something people are gaming right now,

and it is not really showing a benefit.

DR. WILHOIT: I had one technical question about the measure. In the calculation algorithm, which is 2.a.21 on page 9, it talks about if the postoperative glucose is missing either on day one or day two. It says it is a measure category assignment of X and will be rejected, stop processing.

I don't know if that means it is a numerator failure or that you don't even bother to look at it, if there is not a value. So I was just curious, because if the members being -- or the patient is being excluded from the measure because you are missing a glucose, that really seems to miss the point. But I wasn't sure if that was what was meant.

DR. HALPERN: I would find it unusual that any CABG patient would not have a blood glucose the next morning. I mean, they all get labs.

DR. STAFFORD: The other question I have is: In the denominator exclusions, why

would you exclude patients who expire perioperatively? They may have died as a result of their sternal wound infection, because their blood glucose wasn't controlled. So why would you exclude those patients?

CHAIR MORRIS: Any other issues? Okay. Dale, you were present -- Correct me if I am wrong. I think you were present for the time that this measure was initially developed several years ago. So you probably have sort of a --

DR. BRATZLER: I have lived with this measure from the outset. So, actually, you know, the comments that I am hearing actually make me pretty happy when I am hearing that there are a lot more patients getting insulin infusions perioperatively in cardiac patients, particularly on pressures that are driving their sugars up, and other things, because of the known association of hyperglycemia with higher infection rates and higher mortality in cardiac surgery patients,

and indeed, as I was telling some of the folks in the room, increasing evidence that hyperglycemia is a risk factor for infection in many other operations also.

A couple of points really quickly: The measure is not about intensive insulin therapy. I have pushed back on that many times before. We have never pushed anybody to drive down to 110. We always set the control limit at 200, and the current national recommendation from the American Society of Clinical -- or the American -- the clinical endocrinologists and ADA now are, for hospitalized patients, 140 to 180 is the recommended range, and I think that is quite reasonable, and we are more liberal than the national recommendations.

The third thing that I do agree with is that 6:00 a.m. blood sugar is arbitrary, and that is by design. When we were initially starting the measure, we worked closely with Tony Fenari and his group out of

Portland who had implemented insulin protocols for cardiac surgery for sometime, and we thought about how do we capture the glucose postoperatively in patients who have had surgery.

Now lots of people have suggested all sorts of great ideas: Let's take the average glucose over a 24-hour period; let's look at the proportion of glucoses that are less than a certain value, or other things. But in reality, think about the data collection burden to do any of those things.

So we had to make a compromise here, and that was we could try to have a hospital capture a bunch of glucoses, calculate and then have an algorithm calculate an average, or look at a proportion or other things, or pick one time a day that we would look at just to see if the sugar was 200 or less in that time frame. That is what we did for data collection burden.

There was simply no other easy way

to capture the data on relative blood sugar control. Is it perfect? No. Has it improved a lot over time? Yes.

Finally, the number 200: Is it arbitrary? Well, it was based on the study that was published by Latham and his colleagues out of Vanderbilt that looked at 1,000 consecutive cardiac surgery patients, and they used the cutoff of 200, finding that patients who had blood sugars that were above 200 in the two days postoperatively were about three times more likely to have surgical site infections versus those patients whose blood sugars were kept less than 200.

We wanted to be liberal with our number, because we weren't trying to drive hypoglycemia, but we did feel that 200 was a reasonable number based on Latham's study, and that is how the number was chosen.

Some people have argued that we should use the 140 to 180 range. That is now the current national recommendation from the

clinical endocrinologists, but we've stuck with 200 at this point.

The missing data policy -- Maybe somebody on the phone can assure me. I believe the case is rejected from the clinical warehouse. It is sent back to the hospital to fill in the data point. So they either have to list the data.

Is Wanda or Tory or somebody on the call for the missing data?

DR. JOHNSON: This is Wanda. That is correct, Dale. Rejects from the warehouse only doesn't exclude it from the measure.

DR. WILHOIT: So then just to help me understand, so if it is sent back, if it is rejected initially, it would still come back into the warehouse, but would need that value added.

DR. BRATZLER: Yes. The hospital has to complete their data collection.

DR. WILHOIT: So then if somebody genuinely didn't have a blood sugar done, it

would be a deficient event as opposed to an exclusion?

DR. BRATZLER: That is correct. I believe those cases fail the measure if they don't have the blood sugar collected. So they can't send in the chart and just leave the data field blank. If they leave it blank, the case gets rejected and goes back to the hospital to complete the data point.

DR. CIMA: Could you make the measure 24 hours as opposed to 6:00 a.m. from the time the patient is closed or something, much like we do with SCIP. It is not an undue data burden to do that.

DR. BRATZLER: Yes, I suppose we could think about whether there is a way to look at a set period of time, you know, the closest blood sugar 12 hours after closure or 24 hours or whatever the time frame. You know, 6:00 a.m. is what we chose, just to have an arbitrary once a day time so the abstractors would be able to look at the chart

at one point in time and take a look.

I mean, ideally, you know, as I mentioned, we would look at total glucose control. I agree that, if I am in the hospital setting, that is what I want to do, but for measurement purposes to keep burden low, that is what we did for this particular measure.

DR. DUTTON: I don't think that would be an undue increase in data burden now, because science has marched on since this measure was first created when the 6:00 a.m. was the glucose that went to the lab, and that was the one that was in the system, and it was easy to get. But now I suspect that most of us are measuring it hourly using wireless devices that put all of that in the computer anyway, and picking out any one is no harder than picking out any other one.

DR. MORTON: The other thing about data burden is that the person who usually gets this particular measure gets the other

SCIP measures as well, and one of them is, you know, within 24 hours antibiotics are discontinued.

So I don't think there is going to be anymore data burden around that, and the 6:00 a.m. thing is just -- As mentioned before, there's cases that go pretty late, and you've got very little time to kind of get that blood glucose in order.

CHAIR MORRIS: Okay. Were you going to say anything about the VPS with regard to this measure at all, the payment system?

DR. BAUS: It is in the VBP proposed rule.

CHAIR MORRIS: Could you repeat that?

DR. BAUS: It is in the value based purchasing Notice of Proposed Rulemaking that is out for public comment right now.

CHAIR MORRIS: And can you describe to the group what the implications of

that are?

DR. BAUS: Can you repeat the question?

CHAIR MORRIS: Can you describe that a little bit further to the group?

DR. BAUS: I am not the VBP person from CMS. I am the measures person. But basically, the measure will be calculated as a composite. Somewhat of a composite of process measures will be weighted as a total.

This is how it is all proposed in the rule. The HCAHPS will be weighted as a total. This is how it is all proposed in the rule. The HCAHPS will be weighted as a total.

So based on the weights of the different measure domains, that is how the hospital is scored. So individual measures, I am not sure as to how their performance will affect the overall score. That is something I would have to get back to you on, but just to make it clear that this measure is, in fact, proposed for value based purchasing.

CHAIR MORRIS: Thank you.

DR. CIMA: That is an important point, because everything else has been based on at least some scientific merit. Their very comment was, multiple times, it is an arbitrary time point. If you are going to do that, then you better have some good science to support it.

DR. BRATZLER: I have got lots of arbitrary things. So, you know, most experts don't think antibiotics should continue beyond closure of the wound, but we arbitrarily picked 24 hours as a measurement point. So I think you do certain arbitrary things in measurement for data collection burden and consistency of the abstractors doing the work.

I mean, I am more than happy to take back the concept of picking a time frame, you know, a set number of hours. I think that is a reasonable thing to consider, but there are lots of things that are arbitrary.

Thirty days is arbitrary for

surgical site infections, but sometimes they happen on day 31. But we do that for measurement purposes.

CHAIR MORRIS: That is why it is important to continue to examine these things and determine when arbitrariness should be mitigated.

DR. KLEINPELL: When you look at clinical feasibility, 24 hours is a much clinically reasonable timeline than possibly 6:00 a.m. for a patient who just comes back at nine at night.

DR. BURSTIN: There is a new STS guideline that just came out in 2009 on postoperative glucose control with very good recommendations, grading all these things. Again, 110 to 180 is the number they have put in here.

Just one final comment. I think we sometimes get confused about a guideline versus a measure. So the guideline is more clear. The measures -- some of these are

truly just expediency of being able to collect the data consistently across all hospitals in America.

So I think the issue is when does the science, in fact, make that decision for expediency not work. I think that is really the issue that we have given to Dale to consider and bring back to us.

DR. CIMA: That 6:00 a.m. number is not a hard and fast. Not everyone is sent in at six. It could be 2:00 a.m., the most closest one to it, which could be the first blood glucose for a guy that got up at 11:00 a.m. and midnight. So that is the main concern, is that it is not designed, as the other ones, although arbitrary, we are more reasonable in their clinical attempt to say 24 hours as opposed to 6:00 a.m.

CHAIR MORRIS: Thank you. Are there any other comments before we move on to the vote? Okay.

So first: Does the measure meet

NQF criteria for importance to measure and report? Sixteen say yes; five, no.

Next vote: Does the measure meet NQF criteria for scientific acceptability of measure properties? Two say completely; 12 say partially; 7 minimally.

Next vote: Does the measure meet NQF criteria for usability? Five say completely; 6, partially; 10 say minimally.

Does the measure meet NQF criteria for feasibility? Five say completely; 9 partially; 7 say minimally.

Then lastly: Does the measure meet all of the NQF criteria for endorsement?

Before we vote, the major issues that were raised were the sense among the Steering Committee that there is a need for more flexibility in this measurement to better look at the global care to apply to a variety of patient situations or times of departure from the operating room or differing times of closure; and then also a concern about the

possibility of unintended consequences, specifically hypoglycemia.

I think that that was clarified by Dale, that the measure was staying at 200 in order to avoid that. Of course, there probably will be more events of hypoglycemia. I don't think we have any hard numbers, but it is certainly a risk.

Anybody want to add to that at all? Dr. Cima? Okay.

So does the measure meet all of the NQF criteria for endorsement? Nine said yes; 10 said no; two abstained.

This is tricky, because it is very close to a tie, and I think that we should probably revisit this as a Steering Committee, ask for you guys to review this and think about changing the flexibility and the timing of the measurement, and then bring it back to us. Anybody disagree with that?

Allan, do you want to add anything?

DR. MORTON: I was going to say, I think that is exactly it. We all agree this is a laudable goal to get blood sugar better. The number is set at a rate where hypoglycemia would be relatively rare.

The only quibble we have is just the logistics about doing this, because surgery has become 24 hours, and the 6:00 a.m. time frame is not one that is, I think, measuring what we really want to get at, and the within 24 hours would get at it without an undue burden, because the data abstractors are doing the same thing already for other SCIP measures.

DR. BURSTIN: Let's just let Dale and CMS respond to the concerns of the committee, and then we will re-vote and reconsider after that point.

DR. BRATZLER: So, I mean, it is a little bit tough to respond immediately without going to -- there is a technical panel, an expert panel that does meet

periodically and discuss this performance measure.

So it is tough for me to speak for that technical panel, but I think it is a reasonable request to go back and ask about changing the time frame for the collection of those two glucoses, those postoperative blood sugars, and I don't see any big problem with that. I just can't make that statement at the moment without going to the technical expert panel.

There are individuals that we task to actually periodically review these measures.

CHAIR MORRIS: Understood. We are going to go out of order for the next measure and ask for Dr. Cima first to introduce 0218, surgery patients who received appropriate venous thromboembolism prophylaxis within 24 hours prior to surgery --

DR. ROGERS: Could I ask a question before we do that? Terry here. Is

our task the next time we meet, in fact, to revisit some of the issues that have been questionable or had some discussion at this meeting?

MS. MURPHY: The next time that the group meets in person, it will be to look at the next group of measures. What we will be doing is to set up some conference calls to talk about some of these issues between now and that time.

CHAIR MORRIS: So measure 0218, and then we will move on to Dr. Zambricki.

DR. CIMA: This is measure 0218. As already pointed out, it is the number of -- It is a measure to assess patients who are getting appropriately ordered VTE prophylaxis administered within 24 hours prior to surgery or the 24 hours after surgery end time.

This is a continuing measure.

CHAIR MORRIS: Let me just interrupt you for a second. This is not the patients for whom it was ordered, but rather

those who received it. Right?

DR. CIMA: Yes. Oh, excuse me. Received it, yes. Sorry. I as thinking about the other one -- who received appropriate veno-thrombo prophylaxis 24 hours prior to or 24 hours after surgery.

This is, like I said, a continuing measure. The overall goal of this measure is to ensure that patients -- any patient, basically, who is hospitalized is considered a high risk patient for veno thromboembolism, and that we want to ensure that, although things may be ordered, as the other measure is, that they actually are documented as being performed, at least for the first 24 hours or, in some cases, actually administered before the patient enters into the surgical suite.

The rationale behind this is clear. VTE is a major morbidity of patients. A recent Enox study, which was discussed, the number one cause of 30-day mortality in cancer patients after surgery is related to veno

thromboembolism, one of the critical events, which is pulmonary embolism. This is an attempt to minimize that risk in these patients.

There wasn't a lot of discussion about the need for this measure in our work group. Everyone agreed that it is a tragic event, if someone has this, and that anything which should be done should minimize it.

There is a lot of evidence to support this. There are, certainly, high risk surgical patients, pelvic surgery, GYN surgery, orthopedic surgery to some extent, and so there is a lot of data out there. There is also a significant number of trials that have looked at different interventions, and these are all documented well in here.

The numerator and denominator are pretty clear. It is basically those patients that are having these surgical procedures, a very sort of broad spectrum, major abdominal surgery, GYN surgeries, orthopedic, total

knees and hips, cardiac surgery, and sort of the whole gamut of major surgical procedures.

The exclusions are pretty clear: Patients that have a purely laparoscopic procedure, patients that have a surgery less than 30 minutes, patients who don't stay in hospital greater than 24 hours. Those patients are all excluded for very reasonable reasons.

The data does show a gap, although it is much better now. So that was very heartening, but since it is such a significant morbidity, unlike when we were talking about mediastinal infections where it is such a small number of patients, but a more tragic outcome in these patients. It is a huge number of patients at risk. So there is a big difference between 90 percent and 92 percent, even in the just total numbers. So trying to get to 100 is reasonable.

The only real discussion that we had was almost all of the criteria are based

upon the American Academy of Chest Physicians criteria, which most people agree with, are sort of the gold standards for sort of treatment. However, increasingly now, there is some new data and, particularly, by certain societies, namely, the American Academy of Orthopedic Surgeons, which have made recommendations to their members that use different guidelines, so that the combination of anti-platelet therapy plus mechanical devices is a reasonable alternative.

That would not meet the criteria used for this measure, because that is not in the Chest Physician guideline. So we do -- That was the one issue that was brought up in our discussion, as well as in the discussion of the other measure, which is what is the appropriate order any thromboembolism issues that certain very large clinical societies have recommendations that differ than this one?

I don't know if you really want to

call it a harmonization issue, just a difference of opinion about the science. So that would be, clearly -- I think that is a worthy discussion here. I don't know if it is in our scope to address that.

Other than that, it was very clear. It has been used. It is associated with -- It is in the bundle for value based purchasing. There is no mention about disparities in it.

So that was it. Our work group felt it was supported with that one caveat about what constitutes reasonable prophylaxis in a certain subpopulation where the experts in that field feel differently?

CHAIR MORRIS: Thank you. Dr. Carpenter, can you talk a little bit more about this?

DR. CARPENTER: Sure. Thanks. I think this is, obviously, an important guideline, and I think it is important to have this in here. The question is what is

appropriate prophylaxis, and what guidelines should be followed to satisfy this criteria?

The main difference between the guidelines that the American Academy of Orthopedic Surgeons has published and the Chest Physician guidelines has to do with whether we are trying to prevent DVT or symptomatic PE.

So it uses a different subset of the literature, and the problem with symptomatic PE is it is not as common. So the literature is not as powerful. So the Chest Physician guidelines does have a better level of evidence, but it is designed for DVT prophylaxis rather than symptomatic PE prophylaxis.

The feeling has been these are guidelines designed to balance the risk between clotting and bleeding. The risk for bleeding in certain surgeries is -- The consequences of bleeding are very high. Intracranial procedures, for example, mostly

get a bye on these because of the significance of a bleed postoperatively, and bleeding postoperatively into an orthopedic wound is fairly common because of the amount of exposed bone tissue and other areas in the joint that doesn't close as well.

There is often dead space in these wounds. The consequence of postoperative bleeding into these wounds is very significant. Draining wounds, hematomas, have a higher rate of postoperative infections, and postoperative infections in orthopedic implant cases are very problematic. Usually, it means removing the implant, using a temporary implant, potentially reimplanting the implant later on with, generally, about a 10 percent reinfection rate.

So the consequences are higher. So the focus has been on preventing symptomatic PEs and trying to reduce the bleeding risk. So the guidelines mainly differ in that they -- They are very similar

for most of the things, but they do allow for surgeons to accept a platelet -- anti-platelet therapy along with early mobilization and mechanical prophylaxis as an acceptable prophylaxis, which these guidelines do not.

There is a bit of a work-around with these guidelines, if the wound or the situation is considered high risk for bleeding. So if you consider all your hip patients high risk for bleeding, then they can -- and you document that, then that can be excluded from this measure.

So the differences are significant. They may be subtle, but they are significant, and I think that is really the question. This is an important measure. It is just what guidelines are we going to follow, and the guidelines are under revision consistently.

So, hopefully, over time the guidelines will come closer together, as evidence gets more complete. But those are

the main issues.

CHAIR MORRIS: And, Dr. Cima, can you confirm. So it looks like one of the exclusions is if the provider gives a reason for not administering the medication.

DR. CIMA: Yes. It needs to be documented, but that is one of the exclusions. One of the other work-arounds, if you want to call it, which we know is being done is people giving one milligram of Coumadin and documenting that, which certainly is not therapeutic, but it meets the measurement criteria. So they get one milligram of Coumadin, and then they do other things.

It is well known in the orthopedic community that that is how you work around this.

CHAIR MORRIS: That is interesting. I had not heard of that particular work-around.

DR. CIMA: Oh, yes.

DR. BURSTIN: It is really a work-

around. It is not intended to be therapeutic in any way.

DR. CIMA: No. It is purely a work-around for this very measure.

CHAIR MORRIS: It is every definition of a work-around.

DR. CIMA: Exactly. There is another exclusion, that if you are on Coumadin preop that you are excluded from the measure, because you are anti-coagulated for other reasons. So we have noticed this in our literature, in our review of other practices, that the orthopedic surgeon will prescribe the patient one dose of Coumadin before surgery, document that they were on it, and that is a work-around.

Not that I am criticizing orthopedic surgeons. Some of my best friends are orthopedic surgeons. I am just saying that those outside of my friends do that.

DR. ROGERS: The other comment I would make from my pulmonary critical care

days, what bothers me a little, Dr. Carpenter, is the issue of symptomatic PE, because this is an illness that simply does not give you a clear sign. There is no bumper sticker on the forehead that is saying I have PE.

Oft times, it is missed, set aside as anxiety or whatever. So I understand the protection and the natural protection you would have with respect to trying to sustain and protect your surgical site, but you don't die of a bloody knee, and -- well, you can, but the point is that -- and it may not be pertinent to this conversation, and I am not going to change where our Society's position is. But it is just a little scary from a pulmonary standpoint.

CHAIR MORRIS: Dr. Saigal.

DR. SAIGAL: A question about the documentation. Appendix A that has all the procedures that are being covered -- I don't see where that is in what I received at least, from a urology point of view.

DR. CIMA: Yes. It is not in this, but having spent the last six years with it, basically, urologic procedures, anything that is just a stent, anything that is brief -- prostates are excluded. I am not quite sure why, but they are, but that is in the way it is set up. But it is mainly the big oncologic cases that end of staying, because a lot of the urologic cases are excluded, because they are either purely done endoscopically or they are short stay.

DR. DILLON: Can you just comment on the -- One of the exclusions, at least as I just quickly went through this, is procedures performed entirely by the laparascope. Is that a problem with our surgical oncology patients now, who are all -- I mean, many of our whipples are done laparoscopically.

DR. CIMA: Well, the way the criteria are -- and I can just speak to that, because I supervise our institution's group

that does it -- if any incision is made other than to put the trocars in, then it is considered purely laparoscopic.

So if I do a laparoscopic colectomy and then have to make a 4 centimeter incision to extract the specimen, that is no longer a laparoscopic case. So they are basically referring to diagnostic laparoscopy, you know, gall bladders and things like that.

I think the reason why prostates are excluded is because oftentimes you can bring it out through the port and, therefore, the robotic and laparoscopic prostatectomies are excluded, where open prostatectomies, although they are very -- they are rarer now -- were not excluded.

DR. MORTON: I am not sure if I read it right, but would that mean like, say, laparoscopic gastric bypass is excluded; and, clearly, those patients are at extremely high risk.

We have ignored the exclusion and

continue to give prophylaxis ahead of time, because they are obese. Their BMI is high, obviously, and there is potential for risk. We actually give prophylaxis, even though it is excluded.

I think, you know, with the population getting bigger and bigger, that is something we all have to think about. Cases used to be kind of short and easy to do. With a bigger population, maybe not as much. So those cases used to be short, but not always the case anymore.

DR. CIMA: If it is purely laparoscopic, they are excluded from this measure. Now it doesn't make that it is right, but it is just that is how it is done.

DR. DILLON: I think that is a significant problem then with this, as it is written.

DR. ZAMBRICKI: One comment about exclusions: A perioperative death is listed is an exclusion, if the perioperative death is

due to PE.

DR. WILHOIT: The numerator specification for the measure talks about appropriate VTE prophylaxis, but I couldn't find any definition in the measure itself of what appropriate is.

There was discussion in the background about whether aspirin is adequate or not and the pros and cons and so on, but I couldn't find a clean definition, and it seemed like for comparability across hospitals, it would be very important to have a clear, explicit definition of what appropriate VTE prophylaxis is.

DR. CIMA: In the abstraction details, which are not provided here, they are based almost completely on the ACCP guideline, and it does discuss in some detail what they are in the upper portion.

It is not in the detail that the abstractors have, but it talks about whether or not they should use -- based on these

studies, whether or not it is appropriate to use low molecular weight heparin versus unfractionated heparin versus a combination of both with mechanical.

So those are in the abstraction guidelines. It is not in there. So I don't know if it has to be from a point of view, but it is very clear. The abstractors know very clearly what, for each of the procedures, is required.

DR. WILHOIT: Right, which is a good thing, but I think in the measure itself that we are approving -- you know, this is what goes out to the public, and I should be able to read it and be able -- There should be enough information here that I could go do it and measure and get the same results as an abstractor and, you know, I don't have even the basic information to be able to do that.

DR. CIMA: That is a technical issue. I mean, I know the data is in the abstraction guidelines, but whether it should

be here -- that is up to the Steering Committee. As Melinda has said, we are voting on what we see in front of us and, if it is incomplete, then that should be considered in your vote.

DR. ZAMBRICKI: You know, it seems like 1.c.9 is pretty specific, specific guideline recommendation. They go through each procedure and whether it should be aspirin alone, low molecular weight heparin, etcetera.

DR. WILHOIT: That is saying what the guideline recommends, but it is not what is in the measure. The measure comes under number 2, and the measure itself -- the numerator description does not tell me what to count and what not to count.

DR. BURSTIN: I just pulled up the last ACCP guidelines, and one thing they do specifically note is that for patients undergoing laparoscopic procedures in whom additional VTE risk factors are present, which

I think obesity would certainly count, the guideline developers recommend the use of thrombo prophylaxis.

DR. CARPENTER: These guidelines, I don't believe, follows -- It is mostly ACCP, but not exactly. For example, the INR is not specified. That is why one dose of coumadin might suffice versus a specific INR level which ACCP recommends.

DR. CIMA: The ACCP guideline recommendations do specify an INR to achieve therapeutic effect, but not necessarily in the prophylaxis period. So that is the difference. They do say, you know, molecular weight -- low molecular weight at this weight based dosing is effective at prophylaxis, but for long term treatment you would need, you know, X INR.

CHAIR MORRIS: Any other comments or issues? I would like to just recap the discussion. Of course, we want for you to have an opportunity to respond, but just to

recap: The major points that seemed to come out were that this is very valuable. Everybody agrees with the goals. We believe that they are laudable.

It gave the group pause that these don't harmonize with guidelines from the American Academy of Orthopedic Surgeons, but that was explained in, I think, a pretty reasonable way by Dr. Carpenter, in particular, that the goals are actually slightly different here.

There are issues around laparoscopic surgery not being well defined, and I think that the role of laparoscopic surgery has changed substantially since this measure was first developed.

One of the particular ways that this becomes an issue is, for example, with patients undergoing a laparoscopic bariatric procedure. They are, obviously, higher risk, and they probably should be included in the measure.

In addition, more detail could be more readily available in the measure, and I think this was noted among several measures by the different work groups, that more detail could have been made more easily available, and that would have been appreciated by the Steering Committee, particularly given the very large number of documents that we needed to read to prepare for this.

Then lastly, there is a true -- We brought up gaming the system among several different measures before, and it was something that was more sort of projected, but this sounds like more clearly orthopedic surgeons are gaming the system, probably in their patients' best interests, but we do want to avoid situations where people will clearly game the system in kind of silly ways that are wasteful of resources, time, and a little bit wasteful of our integrity, frankly.

So I wanted to bring those issues up, and I would certainly like to hear your

responses.

DR. BRATZLER: All right. Thanks. It has been a great discussion, and I am just launching at the bit to respond to some of the issues.

So let me start with a couple of issues. A whole lot of things have been raised. So to the question of appropriateness, both of the VTE measures that are submitted, VTE 1 and 2 that we call them or 0217 and 0218 -- both of them use the same specifications for what is recommended prophylaxis, which is based largely on the American College of Chest Physicians' recommendations that were published in 2008, with minor revisions.

The performance measure looks at, basically, the hospital abstracts of what was given to the patient, and then the algorithm calculates performance based on whether or not the forms of prophylaxis given to the patient were consistent with guidelines.

So the hospital abstractor actually doesn't have to know what the guidelines say. They simply abstract what was actually given to the patient, and then the algorithm calculates whether or not it was consistent with the guidelines or not.

There was a lot of conversation about the potential out for passing the measure if the patient has bleeding risk or the issue that we have discussed with our orthopedic colleagues.

The performance measure basically looks at those forms of prophylaxis that are recommended in guidelines, but clearly, we recognize that some patients can't take, for instance, pharmacologic prophylaxis. You can't give a shot of an anti-coagulant to a patient who has had a bleeding ulcer or you are concerned. Maybe they have a low platelet count or other reasons.

When we developed the performance measure, we tried not to try to define what

the list of bleeding risks are, because there are just so many different things that could be considered bleeding risk.

So we leave that completely up to the clinician at the bedside. If they document that they are concerned about bleeding risk in any way, then they can use mechanical prophylaxis on the patient, and the case will pass the performance metric.

We do the same thing for neuroaxial anesthesia, even though neuroaxial anesthesia is not a contraindication to pharmacologic prophylaxis, if neuroaxial anesthesia is used, the case will automatically pass with mechanical prophylaxis, if that is used.

Similarly, if the orthopedic surgeon, as I was telling Dr. Carpenter -- If the orthopedic surgeon is concerned about bleeding risk, they don't want to use something because they are concerned about a wound hematoma, then they can document that,

put mechanical prophylaxis on the case, and the patient will pass.

The reason that we have not incorporated some of the issues around the AAOS guideline I have discussed with Dr. Carpenter and on many national agendas -- a couple of reasons.

Number one, it was mentioned before that the AAOS guideline focuses only on symptomatic pulmonary embolism and did not focus on the literature around DVT, and I think our technical panel was concerned about that, because we know that patients who have DVT may have recurrence of their disease years later, well outside of the surgical time frame, but does put those patients at risk for recurrent DVT and potentially pulmonary embolism in the future.

The second thing is just one problem with the AAOS guidelines. All of their recommendations have Level 3 grade of evidence, and that was a problem; whereas, the

performance measure is based only on the grade 1 recommendations in the ACCP performance measures -- or guidelines.

A couple of issues about laparoscopic surgery: We completely agree with you that most patients having these major laparoscopic operations should get VTE prophylaxis. They should also get antibiotic prophylaxis, when appropriate.

So when we designed the measure, we painstakingly went through the list of ICD-9 codes and tried to only include operations in the denominator for which VTE prophylaxis is routinely recommended.

The only laparoscopic cases that get excluded are those that are done entirely by laparoscope with no other incisions, and that actually came up when we originally got the measure endorsed by NQF, because there was concern about excluding laparoscopic cases.

It turns out that nationally only about one or two percent of our cases get

excluded because of that data element, because we have such a strict definition. If there is hand assist, if incisions are extended in any way, then for data collection purposes the hospital has to say, no, this is not a laparoscopic case, and the case is in the denominator.

In fact, the exclusions are so rare that we are now contemplating simply removing the laparoscope data element, because it is rarely used to exclude cases from any of our measures. So it will make abstraction easier, and it is going to have minimal impact on the measures.

Finally, the issue of gaming is one that our technical panel was very concerned about, because we, too, have heard the concerns about use of single dose prophylaxis to pass the measure. It can happen. So what we are actually considering is another performance metric.

Our technical panel has asked us

to evaluate it. We actually have a learning lab that will be testing it in the near future, looking at continuation of prophylaxis up until the day of discharge or day seven, whichever comes first; because, really, when you look at all the guidelines, they suggest continuing prophylaxis until the patient is discharged from the hospital or for at least a week postoperatively.

There is no published study of DVT prophylaxis that is used less than a week of DVT prophylaxis. So we are addressing that, but we are planning to address gaming through an additional performance measure that we will submit in the future.

CHAIR MORRIS: Dr. Morton, would you like to add to this discussion in terms of bariatric laparoscopic cases, particularly?

DR. MORTON: Yes.

CHAIR MORRIS: Before you start, let me just say one other thing. I think that it is important not just to stop the gaming,

but to look carefully at the reason for the gaming. It is there for a reason. So I think that addressing that might be more fruitful in the long run than simply stopping the gaming.

DR. MORTON: I am still a little confused as to whether or not the laparoscopic cases are excluded. There is probably 150,000 gastric bypasses being done a year. They are almost all laparoscopic now.

If you look at the most recent data, about 90 percent are laparoscopic, and they carry very high risk, and they are all done with just making incisions with a trocar. There is really no extraction for any of these.

So from what I heard, it is that you said very few cases end up making a difference for the denominator, but that is 150,000 cases that should probably be included.

DR. BRATZLER: So Tory or Wanda or whoever is on the call, do we have bariatric

surgery actually in the denominator for the measure at all? Is it on the appendix, the tables?

DR. JOHNSON: I think we are going to have to look real quick to make sure. I do have a feeling that there are a couple of bariatric surgeries, and we will look real quick.

DR. BRATZLER: I don't have the number for bariatric surgery of the exclusion, but I can tell you for the data element laparoscopic, because we are so strict for the hospitals about when they can use that data element and say yes that very few cases nationally, across all operations, get excluded. I can't tell you what the proportion of the bariatric is.

DR. MORTON: Well, it is a real opportunity for quality improvement, because those patients should be getting prophylaxis. I know there is some concern in the bariatric surgery community about staple line bleeds and

issues like that, but that has never been proven through the literature.

So it is a real important segment of the population at target because of the increased risk. Also keep in mind, about half of all the deaths that occur after bariatric surgery are due to PE. The other half is roughly leaks. So it is something that really should be addressed, especially with more and more of these cases being done.

DR. CIMA: I can tell you just from our experience looking at this that they are excluded. Our abstractors do not -- If they are done purely -- We do a lot of revisional ones that are open, but Mike Starry, you know, does a lot of those, but for the straightforward bariatric cases, lap bands and things like that, those are all just basically excluded from the analysis.

Now we have a very rigid VTE prophylaxis in those patients, but as far as the measure goes, they are excluded.

DR. MORTON: I can tell you for a fact, they are excluded at Stanford, too. We still go ahead and give the prophylaxis, though. So I think we are just missing it with the measure where laparoscopic bariatric surgery isn't cover for a high risk population.

DR. CARPENTER: If the measure was just left to patients 24 or less were excluded, would that get rid of most of these laparoscopic procedures that are completely -- that should be excluded anyway, the simplest laparoscopic procedures that could be excluded, and could you just eliminate the laparoscopic exclusion altogether, keep the 24-hour exclusion?

DR. BRATZLER: And that is actually exactly what we are doing. So right now the performance measure is actually -- it is not 24 hours. It is actually any patient who has a length of stay that is less than three calendar days. In other words, if they

are in the hospital for less than two nights, they are excluded from the performance measure, because I am aware of no study that has ever shown that a single dose of prophylaxis in the hospital impacts DVT rates.

So that takes care of many minor operations that are done laparoscopically. You are correct. But our approach right now is that we are in the process of looking at simply removing that data element from the data collection laparoscope, taking it out of the algorithms, and then all of the operations that are in the denominator will stay in the performance measures, because we are excluding so few cases right now.

Again, I can't tell you the bariatric specific numbers, but nationally for all operations, we see about a million operations a year in the dataset. It is a very, very small percentage that get excluded.

MS. ZAMBRICKI: I have two questions. One had to do with the idea of the

exclusion of perioperative death, and I was wondering --

DR. BRATZLER: Yes, and I should have corrected that for the previous conversation about the cardiac surgery also. Perioperative death is defined as in the OR or through the PACU. So there is no chance to give either insulin drips or VTE prophylaxis. So if they die in the immediate perioperative period, they are excluded.

MS. ZAMBRICKI: Then my second question was: It looks like the denominator exclusion is patients who stayed less than or equal to 24 hours postoperatively. You were saying something about three days and two nights.

DR. BRATZLER: Yes. So I can't tell you the exact date. Tory, can you tell me the update? The measure was always supposed to be three calendar days, which is two nights in the hospital. So they may say 24 hours.

DR. CIMA: 2.a.10, it specifically says three days, but elsewhere it says 24 hours, but in the exclusion in 2.a.10, denominator exclusion details, like maybe 75 percent of the way down it says patients with hospital stays less than or equal to three calendar days.

The only issue with that now is with clinical pathways. Most bariatric patients are probably out the door the morning of that third day, if not even the day before. I know 50 percent of our colectomies are out of the hospital on day two.

DR. BRATZLER: That issue has actually come up in the orthopedic world. Dr. Lieberman updated us that there are increasing number of overnight stays for certain joint replacements where there is pretty good evidence that those patients should be continuing prophylaxis in the ambulatory setting.

I don't know in the bariatric

surgery literature, even for somebody that has a one or two-day stay in the hospital, is there good evidence on DVT prophylaxis in that immediate -- for those extremely short stays.

DR. MORTON: No. There is not a lot of good data yet, but we do know that most of the time when there is a clot that is formed, it is generally on the table, because that is when patients become veno-dilated, and that is when the clot forms, and that is where the prophylaxis would make its most benefit.

If they already have a clot after surgery, I agree. That is a different story, and there isn't a lot of consensus about how long to extend it, but a single preoperative prophylactic dose makes a lot of sense.

CHAIR MORRIS: Any other --

DR. JOHNSON: There are gastrectomy codes collected for the VTE measures. And, Dr. Bratzler, the correction for the length of stay will be fixed with the April 11 manual.

DR. BRATZLER: Okay, but you said gastrectomy codes. I understand that, but what about lap, just the banding procedures and others that are purely done laparoscopically? I don't know that those codes are actually in our denominator.

DR. MORTON: So for gastric bypass, it is 4431, 4438, and 4439. I've got those burned in my memory, those procedure codes.

DR. JOHNSON: And those are not included.

DR. BRATZLER: Okay, those are not in the denominator currently.

CHAIR MORRIS: Thank you.

MS. ZAMBRICKI: I was just going to mention this might be in the next conversation. The 217 exclusion criteria is different than the 218, even though the algorithm calculation is the same. So it probably was somewhere lost in passing. The exclusion times are different in 217 and 218.

DR. BRATZLER: I can tell you officially it is supposed to be three calendar days, two nights in the hospital, officially, and that is -- The manual is clear on that beginning for April discharges.

DR. CARPENTER: So I just wanted to say before we move to a vote that orthopedic surgeons are in favor of guidelines and the use of these guidelines, and actually, according to Dale, we are one of the highest compliant groups with this.

DR. BRATZLER: That is correct. The orthopedic surgeons have the highest performance in the nation on this measure.

DR. CARPENTER: So this work-around stuff is a minority of situations, but surgeons do want the option of not having to follow these guidelines for some patients that they think it is too aggressive for and could learn to wound complications.

To do that, they do have to use a bit of a work-around, which is better done

with just calling them high risk for bleeding rather than these other things, but the concern really is with what guidelines are being used to determine compliance.

The hope will be that CMS and ACCP and orthopedic surgeons will come together and have a one acceptable set of guidelines that is useful for this measure.

DR. DILLON: If that is true, what should our expectations be in terms of determining whether we go forward with this request, because if we pass it, are we immediately going to put a segment of surgeons or hospitals at odds or out of compliance with this?

CHAIR MORRIS: I think one of the issues, and potentially one of the reasons that orthopedic surgeons are so overwhelmingly compliant with this measure or adherent to this measure is that they are actually gaming the system. So they look adherent, even though for good reasons, they may not be

adherent to the spirit of the measure.

DR. BRATZLER: I actually don't think that is the case. I don't have the numbers in front of me. We have actually -- We can look at the case level, at the actual use of prophylaxis, and it turns out that, if you just use ACCP recommendations, orthopedic surgeons have the highest rates of performance on this measure.

Most actually do use pharmacologic prophylaxis for their hips and, if they don't, they use mechanical prophylaxis, and there is a way that they can document if they are concerned about bleeding risk.

So I don't think there is -- I think there is some gaming that happens. I don't think it is the majority, and we can look at the actual case level data and see what is actually being used for each type of operation.

DR. BURSTIN: Just one process point. If the guidelines evolve and the

measure changes, we do have an ad hoc review policy. We can bring the measure back in at anytime. It will probably come back to you guys, too.

DR. BRATZLER: Yes, that is the other point I would make. We actually have a technical expert panel. AAOS is represented on that panel. We actually update minor details every three months, and they go into the manual every six months.

So if new guidelines come out that change specifications, we change the performance metrics.

DR. CARPENTER: So I think Melinda said we can pass things with a recommendation, with sort of a tag that says we recommend that these differences be worked out, rather than this is the winner and this is not the winner. The guidelines abstraction do follow the ACCP, not completely, not letter for letter, and it says appropriate guidelines.

So there is, I think, room to

follow the recommendation, or to follow the measure, but tweak the recommendations, the guidelines that are followed even before it is re-reviewed.

DR. DILLON: So there are two key points then,particularly pertaining to laparoscopy as well, that this has to be addressed. So just that our recommendations going forward need to have both points included.

DR. WILHOIT: Thank you. The third thing that I think, when it goes out for public comment and so on, if it passes here, I think the numerator description needs to define what is counted in the numerator, because that does alter how one interprets it, and there just isn't enough detail there to know.

CHAIR MORRIS: Okay. Anything else? Let's go ahead and move on to the vote.

Does the measure meet NQF criteria for importance to measure and report? Now I

will ask everybody to push their button once more, and push Send again. Twenty out of 20 says yes.

Next vote: Does the measure meet NQF criteria for scientific acceptability of measure properties? Six said, yes, it completely meets the criteria; 13, partially; one says minimally.

Does the measure meet NQF criteria for usability? Nine say completely; 11 say partially.

Does the measure meet NQF criteria for feasibility? Thirteen say completely; 7 say partially.

The last vote: Does the measure meet all of the NQF criteria for endorsement?

We had quite a discussion here, and so I am going to make the recap really brief, because I think it has really already been done.

Concerns about gaming the system: There were some concerns. They have been

acknowledged by CMS and the contractors for CMS, but they may not be quite as profound as they initially seemed to be in our discussion.

There are concerns about a need for a better definition of which laparoscopic cases should be included and excluded, or maybe just getting rid of the laparoscopic exclusion altogether, and there is a need for more consistency in language throughout the measure or uniformity of language.

Any other major issues that I am leaving out that anybody wants to bring up?

Okay. Let's move on to the vote. Does the measure meet all of the NQF criteria for endorsement? Sixteen say yes; 3 say no; 1 abstains.

Now I would like to move on to the last measure, 0217, surgery patients with recommended venous thromboembolism prophylaxis ordered, and that is Ms. Zambricki.

MS. ZAMBRICKI: Yes. This measure is surgery patients with recommended venous

thromboembolism prophylaxis. I think all the discussion of the previous measure, 0218, is really the discussion of this measure.

The only remaining issue is the uniformity of language in terms of exclusions in the denominator. Other than that, i don't think that there is anything new to cover in this measure. This is the actual ordering versus the administration.

CHAIR MORRIS: That might be the major thing to cover, and can you describe that discussion in the work group about whether this measure would actually be necessary, given that the other measure is present?

MS. ZAMBRICKI: Actually, our work group on our phone call, we didn't really discuss that.

CHAIR MORRIS: Okay. Well, let's discuss it now. What is your opinion?

MS. ZAMBRICKI: My opinion is that it is not. It is superseded by the actual

event. The compliance was 94-95 percent with the ordering. So it seems that the actual administration would be the relevant measure.

CHAIR MORRIS: Anybody differ with that? I guess our burning question is why have two measures?

DR. BURSTIN: One point of clarification is part of the recent NQF Evidence Task Force report, we very clearly said we wanted process measures to be as close to the outcome as possible, and ones that are more distal that are really covered well by the proximate one of administration should really supersede, and really no need for both.

DR. BRATZLER: I am trying not to get in trouble with my colleague on the left here, so being quite cautious about what I say. We have had some of the same thoughts.

So when we first started these two measures nationally, the performance rates in 2005, we sampled 19,000 Medicare patients, and the performance rate on the measures was 70

percent.

So I am really happy to see that we have seen substantial improvement ranging in the 92 percent range for the measures, with minimal racial disparities, by the way, only about three percent disparity rate for all races.

We internally have been having a conversation about whether it makes sense to continue both of these measures. One is whether the recommended forms of prophylaxis are ordered, and then the second measure looks at the timeliness, specifically focusing on whether it is given in that perioperative period, either before surgery or sometimes it is appropriate to wait until after surgery, depending on the type of surgery and anesthesia.

So they do overlap a lot, and the measures are quite similar. Quite frankly, in our conversations we have been discussing about whether we should move to two measures,

but one that focuses on the appropriateness in timing initially, and then the second one which I discussed earlier about, you know, that would be a new measure submission, would be to look at continuation postoperatively to make sure that patients really are getting effective prophylaxis for their operation beyond just the immediate stay.

So none of -- Again, we have a technical panel that meets this month that will be reviewing some of those issues, and it takes time to test new measures, but we have had that conversation also.

CHAIR MORRIS: So my synopsis of your answer to the question, why have two measures, would be -- and I would like for you to correct me if I am wrong -- would be that you have -- There are two separate measures, because compliance with this was so poor when it was originally developed.

DR. BRATZLER: Yes. So, really, when we started, it was first -- It

recommended form of prophylaxis ordered for the patient. That was the first issue. then the second one was timing appropriate. Were they giving it in that close perioperative period? So that was how we saw the difference between the two, was recommended form of prophylaxis, and was timing appropriate.

CHAIR MORRIS: Thank you. Anybody want to say anything else about this measure? Let's go ahead and move on to the vote.

Does the measure meet NQF criteria for importance to measure and report, and specifically around impact, a performance gap, and outcome or evidence? Two say yes; 17 say no. So that means no further discussion of -- or no further voting on the criteria for this measure.

Anybody want to say anything else about that measure before we move on? Dale, would you like to say anything else about it?

DR. BRATZLER: I don't think there is much else to say.

DR. CIMA: What does that mean, though? Now that we have voted no on that, what does that mean?

CHAIR MORRIS: Well, it is not important enough to be assessed as a measure.

DR. CIMA: But in reality, that is one of the SCIP measures. Does that mean it goes away? What does that mean?

DR. BURSTIN: It means that at this point, importance to measure and report is a must pass criterion for NQF endorsement, and you have all just decided it didn't pass the must pass criterion.

So, technically, at this point, unless we hear discussion and follow-up from CMS and Dale that may convince you otherwise to reconsider it, at this point it would be put forward for public comment as not recommended by the Steering Committee.

It doesn't mean it is not endorsed. There is still a long process beyond this meeting, but that at least begins

that process with your recommendation that it not be recommended for endorsement.

DR. BRATZLER: And then I will just make a couple of other points, and Christine can correct whatever I say incorrectly. But typically, NQF has given some grace period. Some of these measures are in -- well, not in the proposed rule about value based purchasing, but the bigger issue that some of these measures are required currently for the Hospital Inpatient Quality Data Reporting Program. I always get that acronym wrong.

So it does take some time for measures to be backed out of the system, but if at the end of the process this measures loses endorsement, then we will begin the process, working with CMS and Joint Commission and others, to pull it out of the measure set for the future.

DR. BURSTIN: So, for example, NQF did not continue endorsement for

the smoking cessation measures in hospitals. That had become essentially check-box measures, not valid indicators of smoking cessation. Again, CMS has continued to use them in this period of time, but they now know going forward those are not recommended for use.

CHAIR MORRIS: All right. Thanks, everybody. We are going to have a moment for NQF member and public comment. I particularly want to encourage those on the phone to comment, if they would like to.

Anybody want to add anything else to our discussion from this morning? Dale?

DR. BRATZLER: I am going to make a member comment that I will make to every NQF Steering Committee, and that is simply about the issue of topped out measures, and Helen knows. She has heard me say this many times before.

Sometimes measures do become topped out, because scientifically valid, good

measures become topped out, because of incentive programs or other things, and I will again make my plea that I am not convinced that we will maintain performance if measures are withdrawn, and if at least there is some way in the future to have a category of measures that are scientifically valid that can be pulled off the shelf down the road, even though -- That is where I worry about losing endorsement for scientifically valid measures. If NQF can figure out a way to have some category of measures that can be resurrected in the future without perhaps having to go through the entire reendorsement process, when they were scientifically valid. They are just topped out.

DR. BURSTIN: That is something we are actually actively engaged in discussing. We will have a discussion with our CSAC this month, actually the end of the month, to specifically see if there is -- it would be interesting to get your perspectives on it --

a set of criteria that you would say no one doubts that this is a valid indicator, a valid reliable indicator of quality. It is just topped out.

Should it be on the front burner of public reporting or should it be somehow put into the background of saying this is a measure that maybe periodically comes up for surveillance, especially if it can be done in a way without a lot of burden, so we don't have to crack a chart to get that piece of information. You can make it more of an electronic surveillance perhaps. Is that something that should remain as sort of some -- we haven't figured out the right word for it yet, but we are working on it.

DR. KLEINPELL: Arden, can I just make a general comment. This is more -- Maybe it is more for the measure, the steward measures. I notice in reviewing the measures that the scientific evidence references oftentimes were 1999, 2002, 2004, and I feel

that, if a maintenance measure is coming forward for review, that the references should definitely be updated.

I don't know if it is optional for them to do that, but that was just a side comment that I had in terms of the measures for maintenance.

CHAIR MORRIS: Thank you for making that point.

Any other issues that anybody wants to bring up? So now it is time for our lunch break, which will be from 12:00 to 12:30. I think that is going to be basically the same as yesterday. I will see you again at 12:30.

(Whereupon, the foregoing matter went off the record at 11:58 a.m.)

- - -

A-F-T-E-R-N-O-O-N S-E-S-S-I-O-N

12:33 p.m.

CHAIR MORRIS: We are going to go ahead and get started here. Our next topic is related and competing measures, and this is really an opportunity for us to go through and discuss, sort of get an overview of the related and competing measures.

I think, hopefully, you guys have this list of related and competing measures that are side by side in two columns, and it basically displays each of the measures that were considered related or competing by the NQF staff.

The goal here in our discussion is just to go through, look at what they are, but not to have an in depth discussion necessarily. We will save that for our next phone conference.

One of the things that we will be doing as a group with measures that we believe are related or that we agree are related are

to ask developers, particularly if it is a single developer, whether they would like to combine these measures or whether they are able or would be willing to harmonize the measures. So those are the sorts of things we want to keep in mind with this discussion.

You can see, so we are just basically getting kind of the bird's eye view here, making comments that you feel are important to bring up at this time, knowing that we are going to have a more in depth discussion later.

So first of all is a cardiac measure, internal mammary artery. You can see the first two -- or the second and third column there, maintenance measure 0134 and measure 0516. The particular difference here, I believe, is that the level of measurement or analysis, which is on the third page, page 3 at the top -- level of measurement analysis in the first column is facility, in the second column is individual. Those are the biggest

difference that strike me. Melinda, are there any other differences that you would like to point out?

MS. MURPHY: No, not that there might not be some other differences within the specifications, but those were the key differences of note from the standpoint of the developer.

CHAIR MORRIS: Okay. The next one is another cardiac surgery measure, and this is maintenance measure 0113 and measure 0456. Participation in a systematic database for cardiac surgery is 113.

Participation in a systematic national database for general thoracic surgery is 456, and this is one where I think that we are probably going to have a particularly interesting discussion.

Again, there is a new generic measure that will be forthcoming, and that will be -- This list will be updated. That will be added. Helen, would you like to add

anything about that?

DR. BURSTIN: Just to point out that I think, as I mentioned yesterday, it would cover all disciplines as opposed to being very specialty specific. So something for you to consider. And I think the issue around does it drive people to use registries in the way we discussed yesterday, I think, is something we need to talk about.

DR. CARPENTER: Is that what was sent out by email yesterday? Yes, okay.

DR. CIMA: The one question I have is, when you say that, though, how is this applied? So let's say your institution participates in X registry. Does that give you a pass on everything else? How can I phrase it ?

So let's say cardiac surgeons want to -- Is this only for cardiac surgery or is this for all specialties? So if I have a multi-institutional practice and I participate in the STS, does that cover my general

surgeons, too?

DR. BURSTIN: No. The STS measure, no. The STS measure is pretty clearly about a cardiac --

DR. CIMA: No, but I am talking about that big measure.

DR. BURSTIN: That big measure would cover anything. Of course, yes, it does. It is not specific to a specific discipline.

DR. CIMA: So does that really meet the purpose of driving quality improvement in one specific area?

DR. BURSTIN: I mean, that is the other question. Could it be stratified? I mean, are there ways to approach it without a separate measure that points people to a specific registry, I think, is the question.

DR. HALPERN: I don't remember if the one we sent out last night covered -- I think you are asking individuals versus facilities.

DR. BURSTIN: It is both. It is individuals, groups and hospitals. Yes.

CHAIR MORRIS: So to be continued, I guess.

Esophagectomy: This was 360, esophageal resection mortality rate, and 361, esophageal resection volume. I thought we had a very comprehensive discussion of the relationship between these measures, and these are both from -- The first two, 363, 361, are from AHRQ, and there is another measure, an endorsed measure, survival predictor for esophagectomy which is from Leapfrog.

So we will discuss whether or not we would request of the developers that they combine these measures, whether we think that that is a reasonable thing to do.

DR. BURSTIN: Let me make just one more point. It is kind of unlikely that they would actually -- these are very complex measures -- just combine them, but I think the question would be is there a way that

particularly the AHRQ measure could potentially -- we talked about it yesterday -- move closer toward incorporating the volume in the way that Leapfrog does.

The Leapfrog measure doesn't have clinical risk adjustment. So the issue is really is there a better mousetrap that you can kind of get to by taking the best of both, and that would be a question going forward, but probably not something they could turn on a dime and do in the course of this project, but more so recommend before the next evaluation.

DR. ROGERS: Arden, if I may, on the first three of these it seems the significant difference is -- on the level of measurement, it is facility agency, and the first three add the individual.

Now if we agree, and we may not, that quality improvement is actually justified and important to change behavior, and that comes down to the individual behavior, there

is something important, I think, in recognizing the identification of who actually -- who individually is responsible for what happens. So I see that as one of the differences between these.

I would personally favor that quite strongly, that we include the individual reference. So I just wanted to comment.

CHAIR MORRIS: Thank you for bringing that up, and please continue to keep that in mind, because this should arise, and it will arise. We will be discussing it more, and where that level of -- where we want to put the crowbar in some ways. Do we want for hospitals -- beg pardon, you don't use crowbars? Would the onus be on hospitals to have their physicians comply in a certain way. Should it be among physician groups, etcetera. We will be talking about that more.

DR. HALPERN: Will we get more details on the Leapfrog measure

CHAIR MORRIS: As we are asked to.

Page 14 is the next measure, and this is -- As Melinda pointed out, it is really a moot point, because the JCH measures did not pass the importance criteria. I'm sorry, Ingenix. I apologize. That is 1479.

Let's see. Then we have page 18, venous thromboembolism. 217 went down as well in terms of meeting the importance criteria, wasn't it? The importance or maybe it was the overall.

So that leaves 0218 and a related measure. It is related. It is not under consideration at this time, and that is 0371, covers medical and surgical patients. It has some, to my mind, substantial differences from 0218, but that is something that we will be discussing as a group.

DR. CARPENTER: If measures such as 0217 didn't pass here, that doesn't mean -- It could be reinstated at another time. Does that mean any one that didn't pass here, we don't consider in the next level of discussion

for those purposes?

DR. BURSTIN: Just one comment -- two. Actually, the measure developer could certainly come back to you and say these are the following points you didn't consider, and you could reconsider it. That is one possibility.

The second possibility is we do put out all measures for public comment, both ones you recommend and not recommend. Although it is not very common, we have had a few instances where not recommended measures -- actually, often in the other direction more so, recommended measures -- the public comment is persuasive enough to make the Steering Committee reassess. So you will have another chance to consider those again.

CHAIR MORRIS: Any other discussion on the related and competing measures for now? We will opportunities to readdress these and again to dig down a little bit.

The next thing on our agenda is gaps to be filled to more fully capture an episode of care. There are about 150 endorsed surgical measures right now. You guys, I think -- is this the list that was received by the group by email? No? Okay. But you will be receiving it.

Having considered the measures that w went through yesterday and today, and then also with an eye to the endorsed measures that you will receive a list of, we would like for the entire Steering Committee to think carefully about topic areas in which further measure development would be useful for quality improvement.

Where do we see the serious gaps, based on your expertise or clinical expertise and quality expertise? So we will be tackling this later, but we just wanted to plant the seed and get you to start thinking about where are the gaps? What measures should be brought up that haven't really been brought up thus

far? Melinda, do you want to say anything about that?

DR. BURSTIN: And one particular thing to consider as we move toward, hopefully, having interoperable electronic specifications, thinking about measures that could be built de novo for that system as opposed to what we are doing now, which is often retrofitting measures developed for paper or claims. So you are, somebody mentioned earlier, looking under the lamplight. There is a lot of that going on.

So the question is have you had good clinical data combined with cost data issues, risk data, whatever it is, what would be the measures you would actually want to assess quality and report on it?

MS. MURPHY: In terms of sequencing the two conversations, one about the related and competing and the one about gaps, is between now and the next time we have a face to face meeting we will resolve the

questions about the related and competing measures for this group of measures.

For the gaps, we really can hold this -- we expect to hold this until after our second face to face meeting where you have had an opportunity to see all of the measures you will be evaluating, but we will go ahead and send you the complete list of endorsed surgery measures so you can be thinking about that.

CHAIR MORRIS: We are moving through our agenda so quickly that I am finding this a little bit alarming. Should we be having more of a discussion about these items right now? Okay.

Well, I think this is probably a good time to go through some of the things that came up repeatedly and for us to basically develop a little bit of a list of the things that we thought were very important that came up repeatedly in our discussions, both yesterday and today.

I can kick this off with some sort

of simple ones. One was consistency of language throughout the measure. That was -- and it should be easy to correct. It should be done before we actually receive the measures. So we would ask for the developers to pay special attention to that.

Another one that came up that was also sort of a simple and fairly concrete thing was the time frame. So the time frame that was listed, consistency of the time frame, and whether the -- and some thought to and rationalization of or justification of whether the time frame is an index hospitalization or whether it is a 30-day period or whatever other time frame is used.

Then, Terry, would you mind just reiterating the point that you had about the importance of the JCH measures that did not pass our importance criteria?

DR. ROGERS: Yes. I perhaps was not alone in feeling a little bad for the person who was at the receiving end of most of

our comments yesterday, and I hope that she got the message, and I think the message should come from us, that our criticisms were not in any way directed at the importance of the issue that was in front of us. It is just that their approach didn't seem to hit the mark with what needs to be done.

Personally, I honestly don't know what all JCH does, but one of the things they might do is embrace the notion of how important the issue of transfusions is, and think about -- I hesitate to talk about mandates -- but to at least encourage, if not require, that hospitals have a very structured and reliable and predictable and responsible way of dealing with transfusions, up to and including perhaps having a transfusion specialist.

I think that where it struck me was recognizing that just measuring a hematocrit is a tiny part of whether somebody really needs a transfusion or not. It has to

do with perfusion and oxygenation and, you know, the whole deal.

So somehow if we get the message back to them that we are very supportive of what they are doing, that it just didn't make it the way they had presented it.

MS. MURPHY: And I think that you did that multiple times yesterday. You reinforced that. The suggestion was made to them yesterday about considering a national patient safety goal that would get at the whole topic area of the transfusion issues, and Dr. Stafford reinforced with them before they left yesterday about the potential for doing just as you have suggested.

It turns out that in their reorganization, their performance measures group and their patient safety goals group are under the same umbrella. They had already made a note of going back to have that conversation with them.

So I talked with them before they

left yesterday. I think they were clear that the issue was the structure and the way in which the measures were put together, not the topic area.

DR. ROGERS: Just one other comment. Certainly, it is a patient safety issue, but -- and maybe things have changed in the past 20 years since I have been doing clinical medicine, but I think one of the issues that we as a profession have to address is to get away from the notion that, oh, just give him a couple of units of blood.

I think it is the ordering piece. We allow people to have this privilege of giving blood who may not have any interest in or engagement with responsibility that is attendant upon that, and I think that is not a patient safety issue. That is a physician or ordering behavior issue that I think we have to take responsibility for.

DR. CARPENTER: Let me just comment while we are talking about the

submitted forms. The biggest challenge for some of us was their abstraction criteria or the -- It is usually in the numerator criteria was often a complex list of abstraction instructions, multiple pages even for some of them that was really code and jargon, referring to other documents.

Usually, those documents, I think, were available if you followed it far enough, but you couldn't do that for all of them. Having some simplified language about what meets the criteria for that measure in plain language -- you know, what is acceptable from the record for meeting some of these criteria -- would be a lot more helpful than the long: This is a yes, if yes is no, and go to the next level and the whole algorithm which the abstractors use, isn't very helpful for us.

So putting that out in plain language, a paragraph of that. If they have to include the other part, fine, but having that up front would be very helpful.

MS. MURPHY: And we can pass that information back to them, but the balance for them is meeting the expectation to have their specifications fully articulated versus having some brief form kind of presentation. But what you suggest may be able to do it.

The other thing is that, in talking with some of the developers and some of the NQF staff who look at how the information is imported into the document, is that some of the things they want to be able to convey are not easily imported into the document. So they default to the position of giving you extra pieces of paper.

So we both need to work some at that.

DR. CIMA: I was just going to say, to follow Ruth's point, I went back and looked at ones that were coming up for maintenance. I think a lot of times, when you are doing like a grant renewal, you have to submit recent literature.

It seemed like for some of the -- one of the maintenance ones, it was as if it was the same stuff they gave 12 years ago. So maybe having a section on -- for that group, that this was the background literature we used initially, and since then there have been this, might be something useful.

MS. STEED: Not only the literature, but their data. Some of them did not have updated data, and in fact, Peter was talking about earlier how several times they have said, oh, well, we actually looked at that, and we are changing it anyway, but we are presenting this now.

DR. KLEINPELL: In terms of the literature, I actually had to go and do a literature search, because the references, I felt, were just way outdated, and it wasn't difficult for me to find updated literature. So I think that should be a requirement for them, not just a recommendation.

DR. SIPERSTEIN: But also a clear

summary of how the measure has impacted health care since it was enacted, because there is a lot of that information, and you really kind of had to --I mean, there are a lot of tables and graphs that were cut and pasted in there, but it really didn't address the question in a succinct way in terms of, you know, has this measure been effective at moving the needle since it was implemented.

So I think that would be helpful for us and helpful for the public that reviewing this as well.

DR. CIMA: Just to follow up on that, you know, with the more and more recent data that has come out, if someone were to bring SCIP 1 as it exists currently, there is a huge amount of data, a lot of it out of NSQIP, out of the VA, that says that that individual measure doesn't mean anything, but the more important measure is actually a composite of if you do 80 percent of these things, then you will have that.

It really begs the question about NQF saying, you know, should we look back and say all these measures now -- now people have been implementing them over the last decade and have looked at it, there is now a huge body of literature that says the individual measures may, in and of themselves, although important as a component of care, do not mean anything, really, if -- You know, there may be a very unique exception in these cases, but it is the composite of doing all of them in a timely fashion that is more important.

I don't know how you get that across, but it would be now a very hard case to make that SCIP 1 per se, if it were brought back -- there is a huge literature that now says it really doesn't matter about the exact timing of it, to some extent.

DR. DUTTON: It is also an answer for what you do with the measures that you think have topped out. Maybe they all go into a pile that becomes your maintenance report

card, that they have effectively become a bundle that you just need to keep reviewing at some lower intensity over time.

MS. MURPHY: And the one thing that comes to my mind -- and Helen, I know, can add to this -- is that the maintenance -- the rigor with which maintenance is approached has continued to evolve for NQF over time, and some of the things that we are asking that developers do at this point, they have not yet caught up with.

We are in the second group of measures in the first cycle of this activity.

DR. BURSTIN: I have mentioned a couple of times these task force reports that we have recently done. So we have done one on evidence, one on testing, and one on harmonization.

All three of those guidances went into effect with projects beginning of January 2011, because we had to have the measure developers have an implementation period where

they have already kind of done their work. They can slip into it, but a lot of the issues you guys have raised are in the new submission form.

There are very fair questions for the newer projects about the use and usefulness of the measure in the field, evidence of importance. Actually, one of the discussions it also had is in terms of measure testing. What should be the requirements for measure testing for measures at maintenance? Should there be new testing done beyond the reliability? Were they done when the measure was begun and, if so, is that testing different? Is that testing looking more at issues of how has the measure actually influenced the performance in the field?

Obviously, measurement alone doesn't do that, but you want to at least be able to say that it had some impact. So I think you will see over time, and you guys are right at the cusp of that, that it will get

tougher.

The evidence report -- if you guys would like to see it, we are happy to share it -- very clearly requires the developers to not just give us the grade and the guideline, but to actually give the quality of the evidence, the quantity of the evidence, and any inconsistencies in the evidence as being the really important consideration for a lot of our committees. With inconsistent evidence, it is really hard to have a measure, as discussed in some of these arenas.

So I think we are trying to make this tougher. Maintenance used to be kind of a pass, and I think the reality is, with so many measures, it is time to just -- some of these ones just need to go away.

In the Cardiovascular Committee last week, or two weeks ago, many of the measures we think of as being sort of bread and butter hospital measures of aspirin on arrival and beta blockers after MI. They are

at 98.5, 99 percent performance, and there is an opportunity cost associated with it. if you are doing that and you are not doing something else that may actually be important to get to the gaps discussion. That is exactly where we are trying to go, but it is interesting.

DR. MORTON: I was just going to mention, maybe for maintenance measures we ought to include impact on health care as a criterion.

DR. DILLON: Right, because one of the things we have to be able to encourage with these is an evolution and a maturation in all of these processes. To me, they are still static. You said cut and paste, and it sounds -- you know, so many of them were cut and pasted out of their previous submissions, and so we don't get to see the maturation.

DR. BURSTIN: CMS has some all or none measures they have developed for SCIP and AMI and THF. We are just beginning to start

to see those. So that is, clearly, the direction, I think, all of us want to go. If we are going to measure these things, it should at least be something -- we do all of them.

DR. HALPERN: Although I do think, like we were talking about before, that some of these, like aspirin on arrival, do need maintenance. Just human nature is, if you don't have to do it, you may forget to do it.

CHAIR MORRIS: I think that that is actually a really important point, and I would like to echo that related to Dale's comment. When these are essentially backed out and potentially retired on a shelf, some explicit method for revisiting them. Once they are not sort of required, are they still being done? I think that is important.

DR. HALPERN: The way to judge that is then to ask them, okay, so how did -- like your question before, how did this impact health care? Did it actually -- what did you

say? -- move the needle, because those would, obviously, be more important to maintain.

DR. BURSTIN: Somebody on our cardiovascular committee, Tom Kottke who some of you may know from Minnesota, just did this great back-of-the-envelope calculation and he said, okay, so if we went from 98.5 to 100, we would save, you know, one life out of -- I mean, just the number reality of it was so striking that, I think, we also want to try to be more quantitative as well and saying, okay, if we are this high up, how much more reality could you move that needle, and how much is really just measurement noise. I mean the noise to signal an issue in a lot of these is not as good as we would hope.

CHAIR MORRIS: Is there another?

DR. DUTTON: It may be that we need the Joint Commission to write us a measure for retired measures as a Joint Commission criteria. Now you pick five off of these 30 retired measures for your hospital to

look at, you know, that kind of idea, but using things that have already been defined, have already been tested but topped out, because they were, so that you are getting some sample nationally of those each year.

DR. BURSTIN: Part of what Dale shared with me is he recently gave a talk to a huge group of hospitals, and one of the things they pointed out was that they were still -- they are very anxious about some of these things coming off the front burner and feeling like that they would go down in performance.

We don't really know that, but I had asked about some of the measures CMS has retired like pulse oximetry and ED for patients with pneumonia, and he said what was interesting was the difference there is that measure had just become a vital sign. You can't walk into an ED without having a thing stuck on your finger.

So I think one of the questions we

have to think through is when is a measure topped out, because we have all just worked really hard to make it top out, and when is it actually built into systems that sort of become infallible. That is, I think, what is not always clear to make that decision.

DR. SIPERSTEIN: I just want to comment. I don't know if the term legacy measure makes sense, but there are different reasons why a measure may be, quote, "retired." I mean, we had a nice example today in terms of we had another measure that really supplanted it.

As we mature, we no longer really care about writing the order about whether the VTE prophylaxis was actually done. The first measure is no way scientifically invalid. There is no problem with it. You know, all of the criteria still stand.

The issue is how to flag it as being supplanted by a better measure or a more mature measure versus a measure that, in the

example you gave, in that medicine changes, and it really no longer is a clinically relevant matter to continue to follow. It would not improve quality to continue to follow it.

DR. DILLON: I just have a quick question for the NQF then, because I was impressed with the composite -- or at least some of the data and literature that we have gotten on the composite scores coming out of the STS. Is this, to me, the evolution of composite scores which will incorporate a lot of processes, like now all of a sudden it is standard of care to get your pulse ox. Is the NQF encouraging the development of composite scores like that to address these issues?

DR. BURSTIN: Yes, and we actually just endorsed the STS composite in our recent Outcomes Committee. We have endorsed several AHRQ safety composites, and I am hoping some of these measures that are out there now, these all or none composites, will come early

as well.

So we are very much encouraging them. I think the issue is it is still an interesting question of which measures should be in a composite, sort of the next step. So one of the issues we had with the cardiovascular composite, for example, that CMS brought forward to us last week is a lot of those measures are pretty close to topped out. So you wind up with, even in a composite, a very small, narrow range.

So something for you guys to help us think through. It is just does that even make sense? If they are really high and you put them together, they are still really high. The all or none helps a bit there.

DR. DILLON: We still struggle with SCIP, though I don't believe necessarily that the individual components, as you said, are that valid, but again the all or none process -- We may be at 98 percent on any given one, but when we look at the all or none

or you start bundling them, then we clearly have room for improvement.

DR. KLEINPELL: This is sort of in a different direction, but what happens -- The woman from CMS said that the 6:00 a.m. glucose was part of a value based performance set that is out for comment.

So we really had issue with that, and we are asking them to come back with different considerations. But could CMS technically move forward based on getting public comment to then say this has to be measured? That seems a little challenging then for clinicians.

DR. BURSTIN: yes. So the process is that the Federal government is obligated to use consensus standards when they are available. They can choose to use -- and, actually, the recent Affordable Care Act made this a bit stronger. They specifically said the Secretary should use endorsed standards and, if they choose not to, they actually

have to post something in the Federal Register and seek public comment.

So doing that, going beyond what is endorsed will be, hopefully, something -- It has got a pretty big burden associated with it, but at times we have had just clear disagreements.

I mean ESRD -- a few years ago our committee refused to put forward an upper limit hemoglobin measure, given all the controversy about EPO, and they are like, no, no, no, it is in the payment rule, we are going forward with this. It is one of those things I was very glad our Steering Committee had actually voted with their conscience, and they were, in fact, correct that there was lots of unintended consequences with going down that path.

It is always an issue for us, just in terms of -- That is why I always just tell committees, just vote with what you think makes the most sense. Ground it in the

evidence.

That is why we are being very vigilant, more so than even four years ago when I came to NQF, that you really are voting on every criteria and subcriteria, because it gives us something to then pass on and say these were the clear issues here. Use, if you need to, but you need to understand what you are potentially choosing to use.

MS. MURPHY: If I heard Dale correctly, the proposal is that this would go into 2013. So my -- Two things. One is that it seems that CMS is better than some about using the NQF endorsed measures and seeking those measures. So they, too, have an opportunity to withdraw from implementation of something that they are talking about two years out.

CHAIR MORRIS: There is another subject that hasn't come up yet in this particular section of our discussion, and that

is the attention to disparities in care. Isn't that one of the core values or core parts of the mission of NQF, is equity in care?

It seems to me that it really got short shrift from most of the measure developers. A few of them cited numbers. For example, Dale did, but one of the recurrent refrains from STS was that they hadn't done it; they could do it, they hadn't done it.

I think that, if this is something that is truly important to NQF, that that needs to be underscored and needs to be attended to.

DR. DUTTON: I would comment on the emphasis that the committee put on outcomes, and the closer you were to an actual clinical outcome, the more we like you in general. I hope that message gets back to the developers, that we killed a lot of process measures that we thought posed undue burden with insufficient evidence that it had any

clinical impact.

That was the biggest problem with the Joint Commission, not that we don't consider transfusion important, but that there was no evidence that their measures would have any impact on clinical practice or on a real patient outcome.

CHAIR MORRIS: I think that that also is a real important point. Previously, it seemed that NQF's focus was really more on processes of care and not so much on outcomes, processes of care because that is something you can actually change, a behavior or an action that can be changed; whereas, outcomes, it was unclear what it is that would have changed the outcome. So a process of care changes the outcome, but were the proper processes of care being identified?

Just as you pointed out, a lot of times the processes that were identified are processes that could be -- in which a change could be measured, really had nothing to do

with the outcomes. It is this ongoing conundrum.

I think that the paired process and outcome measures are probably the most valuable. I want to know if you would like to say anything about that.

DR. BURSTIN: As I mentioned, again this Evidence Task Force report could not have been more clear about the hierarchy of which measures we seek, with outcomes being the highest priority. Process measures with a very strong evidence based link to outcomes is the second priority, so really trying to move down that path.

I think, really, it wasn't so much that NQF had an emphasis on process measures as that the development world was at the process measure stage, with the exception of surgery, actually. Surgery has probably been further ahead on the outcome side and anesthesia than, certainly, most of the medical disciplines have been.

I think that we have started to see that evolution, and it has been very interesting watching as specialties come forward with measures, where they begin, as opposed to -- You know, we are actually seeing some new specialties come on the horizon. They are bringing outcomes to us, and they are kind of skipping all the process stuff we got mired in for so long. I think we need both, when they are good.

CHAIR MORRIS: I do think there is an ongoing problem. Surgical care is so often cross-sectional that looking at outcomes is feasible; whereas, medical care is so often longitudinal that processes are enormously easier to measure. But the truth is that, if we want to change an outcome -- So many times we have said, well, we really can't change this outcome. Then somebody would change something, and they would demonstrate that the outcome could indeed be changed.

So there is something that

happens. There is some action, and that action is a process. We just need to figure out what those processes are.

DR. DUTTON: On that last one, I would point out that the central line infections were a perfect example of that, where 15 years ago, we can't fix that. It's, you know, certain patients are just going to get these. We are going to have that rate. But once somebody started looking at the outcome, we all had a target, and then the ways were found to get there. That is one of the reasons I think that outcomes are really important for driving the process.

CHAIR MORRIS: Thanks. Anything that anybody else wants to bring up in terms of really important points that came up repeatedly?

One, I was sort of waiting for you, Dr. Cima, to bring up was around participation in a registry, what that means for hospitals. What are the implications? If

you would rather not bring this up and just bring it up in the context of our upcoming discussion about related measures, then that is fine, but if there is anything else you would like to say -- Do you know what I am talking about?

It was participation in the STS -- almost mandated participation in the STS registry versus some other registry. Sorry if I am not being clear.

DR. CIMA: The issue is -- and I am still trying to understand who this data steward is from an NQF point of view. I guess cardiac is different, because they already have the market share. They have 95 percent of the market share, and everyone -- as Wanda said, well, everything is there, if someone else can do it, you know. Well, that is a burden.

If you don't really understand the risk modeling and everything, that is a huge burden, especially for the person often at the

institution who is tasked to do this. It is often someone who is not a statistician, is not familiar with it.

So I understand the rationale behind places going to someone and saying, well, we have most of the data and using it, and I support registries, but at the same time, you know, you are selecting -- That person's group has the right answer, and that is the way of doing it.

It just seems that that is not a very open way of doing this. If you wanted people to say what is our mediastinal infection after CABG, well, then you ask them to report what is my mediastinal rate after CABG. You don't ask them to go submit their data, that they have to pay, to send it to somebody else to do the analysis and send it back to you.

I mean, it is picking winners, and I just think -- I don't think that is necessarily the best way to do it, and it is

a burden on institutions; and just because the fact that it is out there and free in the sense of the process which STS uses, that doesn't mean it is easy and doable. It is very complex.

It would be just like saying, yeah, you can collect the 35 variables for NSQIP and here is the algorithm you use to risk adjust it and do all this. A lot of hospitals can't do that. They don't have the sophistication to do it, and I think it would be burdensome.

I don't know if going that way is the best way. And you saw, STS has a lock on it. Is that right? What if an institution doesn't want to use STS? What if they want to use NSQIP? Then you have to redo all your data collection, reprocess it to meet their standard. So they set the standard. It is not right.

DR. DILLON: And it is going to be a challenge as databases proliferate, and that

is what we are seeing right now, either at the state level or the society level or the consortium level. So that is going to be a challenge that NQF and others in terms of data collection are going to face, and for the very reason that we can't be in a position where you specify what database you must be in, in order to meet a certain quality criterion.

DR. MORTON: Just two follow-up points. One is I think this issue is going to get worse over time, because there is going to be pressure on hospitals to say, well, you can only go with one of these systems; you can't go with them all.

As we were talking the other day, you know, every specialty wants to have their own database, and there is going to be competition in the hospital, and they are going to have to choose. So when you set up a potential advantage for one, it can create some issues.

One other follow-up point that I

think Richard was making earlier was about moving toward elimination. I've kind of moved toward that in NSQIP where we are trying to move away a little bit from O to E ratios, because what really matters is getting rid of the complication.

We all know the patients are sicker. Okay, great. That doesn't take away the complication, and we still need to work on those cases, too. So I understand. The risk adjustment, in my mind, is just to make clinicians comfortable with the data at the end of the day, because you still have to change some of those practices and see if there are some things that you can eliminate. You know, we have to start thinking that way. Some of these things could go away.

DR. HALPERN: And I think the other thing about tracking like the O to E ratio was in living in it in a VA system, there is, talking to many people around the VA, the sense that,okay, I am a little afraid

to operate on this patient who is a little sicker, because it might affect my O to E ratio, and I don't think that is good for the quality of patient care, and I can tell you, it is a very real phenomenon.

DR. CIMA: Just to go back to the example we started off with, and I can tell you, at Mayo Clinic I sit on our three-site committee. We have a site in Arizona, Florida and Rochester, and at some point it does come down to finance.

Arizona -- I mean Florida recently said, you know, we participate in STS. We now participate in the multi-specialty NSQIP, because the state of Florida has now sort of said everyone is going to do it. I said, we are going to pull out of STS. We are paying abstractors to do both things. We can get cardiac surgery from NSQIP.

So now we are just going to -- Now I am going to go back and say, well, that's good, guys, but guess what? That is what I

talk about an undue burden. We should not be deciding at the local level what you have to do in that sense.

If they are participating in a registry database, then they should be able to give you the data from that, but what we basically are saying, you can't use that data, you have to use the STS data system and, yes, you can provide it, but STS says you have to do 100 percent of your cardiac cases. NSQIP says I have to do 20 percent.

So now I have to go do the 80 percent abstraction on that, have to format it their way. I have to use their coding. That is inappropriate. That is picking a winner, and that is not what we are supposed to do. At least, that is -- As I sit here as an individual, that is not what I -- Give me the science, but I am not picking the winner, and basically you are. NQF is.

DR. BURSTIN: And again, some of this is because these are historical. These

are the only games in town. They are the only ones who have submitted the measures.

I think I -- It was interesting. After our conversation yesterday, I had a conversation offline with Frank Opelka, who is on our CSAC, and in some ways what you really want to get to is we identify what the measures are. Here is the specifications. They meet criteria, and there is some cloud computing that allows you to just submit wherever, but the problem is it is not there yet.

I think that is all -- We have a sense that ultimately all these registries will, hopefully, use harmonized endorsed specifications, and then submit them however they so choose, but I think that it is a great concept. I don't know that we are there yet, and I think we are hopeful that that is the next step.

DR. CIMA: What do we tell the institutions that have to respond to what we

just passed yesterday?

DR. BURSTIN: They don't have to do anything. It is not required for public reporting that you use STS.

DR. CIMA: No, but if it becomes public reporting.

DR. BURSTIN: Well, I think we would tell them that you don't have to, based on at least what we know at this point. You do not have to submit your data to STS. It is burdensome. I agree with you.

You don't have to submit your data to STS to perform those measures, just like the two recent measures we passed from ACS that Bruce Hall had done for CMS, two surgical outcome measures based on NSQIP, and the agreement was that, if CMS took those forward for public reporting, you didn't have to submit them via NSQIP.

CMS would put forward another data platform for you to submit electronically, but you wouldn't have to be a participant in NSQIP

to do that. I think that is where we are going, and the question is how does that proceed over time.

It will be interesting to see. Same model as well with the ACC registry as well for PCI, is that, yes, you can submit to ACC. We have a PCI registry, or they will also potentially have a -- you know, build an alternative platform. But the sampling issue is a really big harmonization issue. It is the same issue we've got with SSIs, frankly, between the ACS measure and HSN, the CDC measure.

This is a -- Again, this is where we look to what the science tells us. We don't have a horse in this unless -- We are just trying to stay very, very evidence based, and the science tells us 20 percent is adequate, like we know for CAPS 30 patients per practice is adequate, and that is what it should be.

Harmonization has now become, I

would argue, probably 50 percent of our work, and it used to be probably five or 10 percent when I came to NQF four years ago. It is where the game is right now, because there are so many competing efforts, as the stakes have gotten higher. Everybody wants to be in the measurement game in a way that we didn't see before.

Do you read about some of this, Steve, from where you sit as a consumer at Consumer Reports?

MR. FINDLAY: We continue to push for more outcomes measures and are very focused on patient engagement and family engagement measures. That is our big push over the next year, joined with many other consumer groups.

DR. SEARS: What role is comparative effectiveness going to play here?

DR. BURSTIN: That is the second question in two weeks on that. It is an interesting question. Not directly, I think,

other than the fact that I think comparative effectiveness provides a broader evidence base that can be used to support measurement. Beyond that, I don't know.

DR. SEARS: Are they subscribing to databases or are they going to create their own? That is the question, because if they are going to create their own and it is run by the Feds, it may be a solution that we are not endorsing a particular database.

CHAIR MORRIS: Allan, did you have something to add?

DR. SIPERSTEIN: Just to follow up a little bit with what Bob said. Each of our institutions spend a huge amount of time and resources submitting very similar bits of data to different organizations, and obviously, it is a big financial and time resource doing a lot of menial work, where it kind of detracts the impact from really focusing on areas that are important or taking on new projects, new creative projects.

So just to -- We all suffer with these multiple competing standards, and it would be nice in an ideal world to have a single entity to which electronic data is uploaded, an then maybe analyzed in various ways. But part of the problem I see with multiple competing organizations -- and I know in John's world there are two very head to head competing bariatric databases that exist, and it is very time and labor intensive to use both, and institutions are picking one or the other.

From a national impact level, you've got this apples and orange comparison. So you really cannot see how bariatric surgery is evolving over time in this country, because it really is difficult to make direct comparisons.

So maybe just a plea to try to move toward a uniform standard. It may not be perfect, but the whole issue of uniformity is going to really increase a lot of the

efficiency in terms of what we do.

DR. HALPERN: Or maybe one data entry place from which others can then extract their data.

DR. CIMA: Yes, that is the big issue, is using the -- and if you are going to go to what you were saying earlier about risk adjustment, yesterday STS said, well, we risk adjust if it is a re-operation. We risk adjust if it is bad. But what if the risk adjustment is different in the different datasets?

So then, you know, what -- Institutions are supposed to say, well, my internal data says this, but when we send it to them, it comes out differently, because their models are different. That is why I would -- You know, I tell the residents and staff we work with , keep it simple.

You want to know what your mediastinal infection rate is? This is sort of what John was saying. That is the

mediastinal infection rate. Doesn't risk adjust it. Let us know what the rate is, and more and more, there is a lot of data coming out that is saying risk adjustment may be -- We are swinging too far the other way, in that we need to be a little bit more cautious in saying, well, they are a high risk patient. If you know what your infection rate is in that group of patients, then you should try and lower it.

DR. HALPERN: I think that also goes to what Terry said yesterday about what is the cause of the bad outcome in individual patients, and how do we learn from that, rather than looking at just an overall mortality rate -- our risk adjusted mortality rate, which may or may not reflect the real issue.

DR. DUTTON: I think part of the emphasis on risk adjustment has to do with an unintended consequence from public reporting. If you only look at the data privately, if

your quality management data is for your own quality management purposes, then you just want it raw rates and trends over time, and you understand what the risks are, so that you don't need that over-adjustment. But when you present it to the public who doesn't understand all of those issues, or you compare between institutions, it becomes much more important.

CHAIR MORRIS: I was going to ask what you have to say about that, because I think it becomes then to educating the public.

MR. FINDLAY: Yes, which is just a huge hill, maybe even 80 degrees. A lot of us are looking forward to a meeting that AHRQ is hosting on the 23rd, which is an invite-only meeting, I think, just to keep it as a sort of a working meeting, on public reporting, tackling these questions.

I don't think there has been a similar meeting where everyone is called together to sit in a room for a day and try to

hash some of this stuff out. So I hope that we are going to get some clarity coming out of that meeting for at least a path forward. I don't think we are going to get a lot of answers, but I think we will get, hopefully, a path forward.

There are six commissioned papers on public reporting of health care quality information, and I would urge you all, obviously interested in this area, to get hold of those when they come out. They should be out probably right around the 23rd.

DR. BURSTIN: I just submitted mine while we have been sitting here on standardization of metrics. So, yes.

MR. FINDLAY: I reviewed one, and it was exceptional. I think that AHRQ went through some steps to identify excellent people like Helen to write these things. So, hopefully, there will be some galvanizing around that and some coordination around that conference and those papers.

MS. STEED: Helen, can you send us -- Can we get a copy of your paper?

DR. BURSTIN: That's AHRQ. As soon as they are done, yes, I will certainly share it with you. Ours is really about the benefits of standardization and where does standardization allow us to go in a way that we can't move if we are kind of still stuck in this sort of fiefdoms of data and the fiefdoms of measures.

MR. FINDLAY; Yes. There is a huge emphasis on standardization and harmonization at this meeting and how that is going to happen for public reporting.

DR. HALPERN: When are these papers going to be coming out?

DR. BURSTIN: I don't know yet, but AHRQ will publish them on the website. So we will send you the link. Yes.

MR. FINDLAY: They will spread it around pretty fast, I think, on the 23rd.

DR. DUTTON: I wanted to comment

quickly on the data collection burden. I mean, it has always been true that we can afford as much quality management as we can pay for, and then you can look at every -- the process of administration of every individual drug and every blood pressure but, obviously, we don't have the money or the resources to do that.

So there is always the decision as a quality manager, how much can we afford, and what can we stand to look at. But the answer around the burden of collecting data: Some of it will be in advancing technology.

The anesthesia registry that we are building is entirely based on passive electronic data without going through a nurse abstractor or eyeballs. I think, as we become more digitized in the future, that is going to be a more viable model.

The other thing that I have seen at a lot of the large institutions I have been with and visited is that, faced with multiple

reporting burdens, the way they are dealing with that is they are creating their own internal repositories that gets all the data in that everybody might need for any purpose into one registry and then writes reports out of that. So they can just hit the STS button once a month, and the STS report goes off, and they hit the NSQIP button, and the NSQIP report goes off. But that involves, obviously, organizing all your data internally so you can connect it.

Incidentally, it is much harder to aggregate at the national level, because we can't collect identifiers right now. So it makes much more sense to aggregate it locally under the current HIPAA approach.

DR. ROGERS: I may make one other -- or ask a question, actually. We have talked for two days about the process of evaluating services that have already been performed. I doubt it is the charge of this committee, but does the NQF spending time

thinking about appropriateness of care and whether the service should have been actually done to begin with?

DR. BURSTIN: Yes. It is a major emphasis. What we have done, we have just almost completed a very large project on imaging efficiency, third rail for sure, so both radiologic and cardiac imaging in particular. We are now beginning to see appropriateness measures and overuse measures coming into pretty much every single project.

For those of you who -- I was mentioning this to Christopher before he left. Those of you who didn't see it, the Washington Post today had an excellent piece on physician ownership of radiation oncology for prostate cancer and sort of potential conflicts, really, really interesting work. MEDPAC is going to come out with a report, etcetera. So lots coming down the road on overuse as well.

DR. DUTTON: The comment about maybe not doing an operation, because you are

concerned about your O to E ratio -- what the anesthesiologist thinks of that remark is, oh, good, maybe you shouldn't do that operation.

DR. HALPERN: The only thing I say about that is, especially as a vascular surgeon where we have very many sick patients, some surgeries are palliative, and they need to be viewed as that.

So if you have some guy whose foot is rotting off and it is causing him a lot of pain, even though he is sick, he still -- you know, it is a palliative procedure.

CHAIR MORRIS: All right. I think that that was a valuable discussion. Let's see. We have an opportunity for NQF member and public comment, and that is actually scheduled for 2:00 p.m. Is it fair to -- Okay.

So if there is anybody on the line who would like to comment now, please feel free. They are just as verbose as throughout the rest of the meeting. It is really quiet

out there.

I think that there are probably a lot of questions about next steps and the timeline for this project, and I would really like for us to talk a little bit more about that.

Several people came up to me during the break and asked about what happens next for us in terms of telephone meetings, what are our goals, and next in person meeting.

MS. MURPHY: And Alexis and Jessica will have to help here, but one thing we will get you out soonest will be a summary of the information from the voting today, so to get you the numbers back with the major issues identified and the decisions you have made. So you can just take a look at that. Be sure we got it right.

Then we will provide you an updated document on the related and competing measures, and at the time we provide that to

you, probably would query you in terms of your availability for a conference call for the purpose of discussing the related and competing measures in more detail, and offering recommendations about going forward with those.

Then the next activity for the face to face meeting that will occur on -- May 4th and 5th, is it? Alexis says yes -- will be that we will put together a similar set of documents that you got for this meeting and get those out to you, and I guess I would plan that we would reconvene the work groups in the way we did before, but saying that out loud, I know what we need to ask, is did you find the work groups useful to you in preparing for the meeting and the discussions? Okay. So you are open to doing the work groups for the next phase. Okay.

So that in broad strokes, I think, are the things that we will be doing between now and the 4th of May.

Alexis or Jessica, other things that you would add to that?

MS. FORMAN: Once we send you the voting results and the conditions for the measures that you would like for us to send back to the measure developers, we will give them about a two-three week deadline to get the responses back to us, and then we will provide that to you all, and we will try to do it before we send you the Phase II measures, so it won't get too confusing.

MS. MURPHY: The other thing that we will be doing, given some of the conversation today -- which by the way, was very useful to us for the next phase and very useful, I think, to NQF overall -- is that we will go back to the developers whose measures we will be looking at in Phase II and say you might want to know that this Steering Committee finds it very important that these things be addressed, and give them an opportunity to get that done before you see

them.

CHAIR MORRIS: Okay. Is this the first NQF Steering Committee meeting that is finished before the actual time?

MS. MURPHY: No, but pretty close.

DR. BURSTIN: I said joking to Melinda earlier, I mean, it is just something about a room full of people who do surgery. It is just kind of moving on through. As a flea myself, we can circle the evidence for an hour before we make a decision. So way to go, surgical team. Does everybody know the term flea? Oh, yes. The last to jump off a dying dog -- that would be me.

CHAIR MORRIS: Thanks again for your time, your effort. Really appreciate it, and everybody's willingness to play well in the sandbox, and also bring forth all of your ideas.

I would encourage anybody who didn't find an opportunity to speak up quite as much as some others to -- Definitely, your

ideas, your thoughts are very much valued by the group. So please feel free to contribute as you see fit.

DR. ROGERS: I have done this personally,. but I would like to publicly compliment Arden on her superb leadership capability.

MS. STEED: And the fact that you had to do it solo.

CHAIR MORRIS: They say a benevolent dictatorship is the most efficient form of government.

(Whereupon, the foregoing matter went off the record at 1;33 p.m.)

- - -

H

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 Maintenance 2010 Steering Committee

Before: Arden Morris, Chair

Date: 03-01-11

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.

Court Reporter