NATIONAL QUALITY FORUM

+ + + + +

ENDOCRINE MEASURE ENDORSEMENT PROJECT
STANDING COMMITTEE

+ + + + +

THURSDAY
FEBRUARY 27, 2014

+ + + + +

The Standing Committee met at the National Quality Forum, 9th Floor Conference Room, 1030 15th Street, N.W., Washington, D.C., at 8:00 a.m., William Golden and James Rosenzweig, Co-Chairs, presiding.

PRESENT:
WILLIAM GOLDEN, MD, Co-Chair
JAMES ROSENZWEIG, MD, Co-Chair
ROBERT BAILEY, MD, Janssen Scientific Affairs
TRACY BREEN, MD, North Shore-LIJ Health System
WILLIAM CURRY, MD, Penn State College of Medicine, American Academy of Family Physicians
JAMES DUDL, MD, Kaiser Permanente
INGRID DUVA, PhD, RN Veterans Health Administration
STARLIN HAYDON-GREATTING, Pharmacy Quality Alliance
ANN KEARNS, MD, PhD, Mayo Clinic
SUE KIRKMAN, University of North Carolina Diabetes Care Center
ANNE LEDDY, MD, American Association of Clinical Endocrinologists
GRACE LEE, MD, Virginia Mason Medical Center
LAURA MAKAROFF, DO, Health Resources Services Administration (HRSA)
ANNA McCOLLISTER-SLIPP, Galileo Analytics
PATRICIA McDERMOTT, RN, Aetna

JANICE MILLER, CRNP, Thomas Jefferson University School of Nursing
CLAUDIA SHWIDE-SLAVIN, American Association of Diabetes Educators
JANET SULLIVAN, MD, Hudson Health Plan
WILLIAM TAYLOR, MD, Beth Israel Deaconess Medical Center, Harvard Medical School

NQF STAFF:
POONAM BAL, Project Analyst
HELEN BURSTIN, MD, Senior Vice President, Performance Measurement
KAREN JOHNSON, Senior Director, Performance Measurement
KAREN PACE, PhD, Senior Director, Performance Measurement
LINDSEY TIGHE, Senior Project Manager, Performance Measurement

ALSO PRESENT:
NONI BODKIN, Centers for Medicare and Medicaid Services (CMS)
KYLE CAMPBELL, FMQAI, CMS*
KEZIAH COOK, Acumen, LLC
DEBORAH DIETZ, Acumen, LLC*
SOEREN MATTKE, RAND Corporation, CMS
ELIZABETH RICKSECKER, Centers for Medicare and Medicaid Services (CMS)
ALMUT WINTERSTEIN, PhD, University of Florida, CMS*

* present by teleconference

TABLE OF CONTENTS

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

(8:06 a.m.)

CO-CHAIR GOLDEN: Good morning, everybody. Thank you for hanging in there yesterday. I hope you had a pleasant evening. I don't know about you but I woke up this morning and was wondering whether my dreams were valid and reliable.

The scary part was if we were a psychiatry panel, you would now be voting.

So we have another busy day. And I think we have a new agenda. Jamie, do you want to make any comments about yesterday? I know you are eating your breakfast. And we need to move along.

Does anyone want to have any comments or questions about yesterday? I talked with a couple of you and you thought things went reasonably well, all things considered. And we appreciate your attentiveness throughout the whole day. I think people were on track and having good

conversations. We actually got some things done. So, it is good.

We go to 519 and there is a new agenda and that is going to be the CMS measure on foot care. And out discussant is Sue Kirkman.

MS. TIGHE: And I will just jump in. Operator, does Deborah Dietz have an open line?

CO-CHAIR GOLDEN: Is she there?

MS. TIGHE: Cathy?

OPERATOR: Yes, ma'am?

MS. TIGHE: Does Deborah Dietz have an open line?

OPERATOR: I'm not showing she has joined at the moment.

MS. TIGHE: Okay. She might be joining in the next couple of minutes. If you could just give her an open line when she calls in.

OPERATOR: Okay.

MS. TIGHE: Thank you.

OPERATOR: You're welcome.

MS. COOK: Hi, everybody. I am Keziah Cook and I am with Acumen, LLC. We are one of the companies working to develop Measure 0159.

One of my colleagues, Deborah Dietz should be joining us shortly. And unfortunately with the reschedule, none of our other colleagues are able to join but Deborah and I should be able to answer your questions.

The measure captures whether there is documentation in the home health clinical record of both patient education and diabetic foot care that includes monitoring of the patient's lower extremities for evidence of skin lesions.

Just to give a little background about why CMS believes this measure is important to continue reporting.

CMS began publicly reporting quality data on the Home Health Compare website in 2003. And initially all of the

reported measures were outcome measures. However, we received significant input from NQF and others that process measures, including diabetic foot care in patient education were important. So, these measures were added in 2010 to capture aspects of care that are directly under the provider's control and to capture quality of care even for patients who aren't likely to improve.

In addition to providing information to consumers, the public reporting of the measures was designed to reduce variation and practice and otherwise to incentivize agencies to adopt the best practices of care for patients with diabetes.

The measure does seem to be moving practices in that direction. There has been steady improvement since adoption in 2010 and CMS believes it is important to continue reporting this measure. However, the measure is based on generally accepted standards of care. These standards of care applicable to

home health. The measure includes both foot care and monitoring and also patient education. So it important to note that even though some of the literature regarding interventions relating exclusively to patient education are a bit mixed, we do also include monitoring and foot care.

Between 2010 and 2013, the home health agency average score on this measure increased from 89.1 percent to 93.4 percent. And this is based on all Medicare home health agencies with at least 20 home health quality episodes.

In 2013, there were about 850 such agencies. That is about 71 percent of all agencies. And as mentioned previously, this is an existing measure that has been publicly reported since 2010.

MEMBER KIRKMAN: Okay, so as you heard, this is the measure is Diabetic Foot Care and Patient Education Implemented. CMS is the measure steward.

The description of the measure, as you heard, is the percentage of home health episodes of care in which diabetic foot care and patient/care giver education were included in the physician-ordered plan of care and implemented for diabetic patients. And the rationale is, as she said, that instruction on foot care is expected to influence health outcomes, fewer acute care visits for EDUs, acute care hospitalizations, through reducing diabetic foot ulcers. There has been improvement in the measure over time.

The numerator is the number of home health episodes. We are at the end of the episode diabetic foot care and education specified in care plan had been implemented. So, my understanding is the numerator is basically a checkbox on the form, although there is significant sort of background documentation that exists. And the denominator is the number of home health episodes of care, ending with a discharge or

transfer to inpatient facility during the reporting period, other than those covered by generic or measure specific exclusions.

And the denominator exclusions are the patient is not diabetic or had bilateral foot and lower leg amputations and also episodes ending in patient death.

It is a process measure. It is collected purely through electronic clinical data and the facility, so the home health agency is the unit of analysis.

In terms of the evidence, the measure developers submitted 88 guidelines regarding foot care and education. They also submitted a systematic review, a Cochrane review of patient education. The conclusion is that there is insufficient evidence showing that limited patient education alone is effective in reducing diabetic foot ulcers. I think sort of like the discussion we had yesterday, I think most of the literature is about a sort of comprehensive path of care

from risk assessment and then referring high-risk people into some sort of comprehensive program that includes foot care and patient education but it is hard to kind of separate out any one component.

The other comment I had about the literature is that I am sure it is all ambulatory care patients. So, it is here being applied to home health care patients.

So, do I stop here and we talk about evidence? Okay.

CO-CHAIR GOLDEN: Okay, was there -- and I guess discussion of the committee group, the subcommittee was reasonable, unreasonable -- the question was by itself, the evidence would be limited for effectiveness as opposed to a component?

MEMBER KIRKMAN: Right. I think the discussion was, and again, our workgroup had the four foot care measures so I am sort of forgetting when the discussion happened. But I think for all of them it was the same

discussion that there is evidence but it is typically not evidence for X component versus not doing X component. It is typically sort of a comprehensive program versus usual care. So, it is very hard to kind of separate out the evidence --

CO-CHAIR GOLDEN: It doesn't necessarily stand alone but is a part of a continuum and so forth.

MEMBER KIRKMAN: As part of a continuum, there seems to be evidence. And again, for this particular measure the patient population studied, I don't think has typically been a home health population. But you could say in many ways they are probably older and have more vascular disease and probably may be higher risk for foot ulcers. So, it sort of makes sense that they would be a high-risk population.

CO-CHAIR GOLDEN: Comments and questions about evidence>

MEMBER SULLIVAN: I have one

question for the developer. How is the foot care education defined for them to be able to check off a box? And also the exam, what is their definition that they have to follow? And how do you know if all of that is followed? I know you do chart reviews.

MS. COOK: So, I can probably pull up the exact guidance in a moment, if you are curious. I believe that the education component specifies that it is a comprehensive education meeting standard practices. The foot care component itself requires physical examination of the feet, looking for lesions.

MEMBER MILLER: Are you talking about diabetes education specific to foot care or is it comprehensive diabetes education?

MS. COOK: It is diabetes education specific to foot care.

MEMBER MILLER: And are there specific components that are identified in that?

MS. COOK: The general approach

with the various home health interventions that are captured in the process of care measures is to leave the specific interventions up to the home health agency and the patient's physician. What CMS collects is whether or not interventions that meet best practices were incorporated into the patient's plan of care and implemented.

So the exact interventions are left to the clinical discussion of the care team.

MEMBER MILLER: Thank you. So, it is kind of similar to what we talked about yesterday with diabetes education being so varied and so non-specific from agency to agency. Thank you.

MEMBER CURRY: Sue, if you said this while I was looking at the evidence algorithm, forgive me. But has your group recommended it is rated as insufficient evidence with exceptions. Is that what you came to?

CO-CHAIR GOLDEN: You may not have gotten there.

MS. COOK: No, we did get there.

MEMBER KIRKMAN: Well, so it is a process measure. There is a systematic review with QQC ratings. And then yes, I think we would get to 5c, that the systematic review concludes that -- let's see.

Yes, so the systematic review concludes that there is insufficient evidence showing that limited patient education alone is effective in reducing foot ulcers. But I guess we felt like again, you have that problem of trying to isolate out one specific thing and also this measure is more than just patient education. It is also inspection of the feet, for whatever that adds.

But yes, I guess if we --

MEMBER SHWIDE-SLAVIN: I was wondering how you got to insufficient evidence with exception because we when we followed the algorithm, we got to raise level from 5c.

MEMBER CURRY: So, one might look at box 3 and then if the answer is no to above, go from that point. Because is the evidence really about the measure or is it about other things around the measure and not directly related to the measure.

Then you could go to box 7 and then you could go to box 10, and then walk over to what is insufficient with exception.

MEMBER KIRKMAN: Yes, although there is a systematic review with grading of the evidence. I'm not sure. I'm sorry.

DR. PACE: And I think the point that you made about the measure is more than just the education. So, if it were just the education, you might up picking the low box. But there still is an option to consider the exception. But I think the question is, with this having both components, is the other one, as you were just describing, get you down to looking at this as an exception.

And I guess the other question is,

is this any different -- you mentioned you were looking at the other or yesterday looked at the other foot measure. Is what is being measured here much different? And would the evidence requirement be different than what you looked at before?

CO-CHAIR GOLDEN: So, just to parse this out, you start to parse out what evidence you are looking for. So, I guess the first question is is any kind of education about foot care a valid activity or just foot care in general? And then you get into the question of does this intervention that they are measuring in the house have evidence of value?

So, I think depending on how you parse that, you can get into different supports for the measure, I would think.

MEMBER KIRKMAN: Right. The other thing is the Cochran review talks about limited patient education. And so again, I don't know whether your education is more than

limited. I mean, it is just very difficult. And it is a lot like the measure yesterday. I am reluctant to say that the evidence is low because I think we are hampered by this, trying to isolate out a specific component of a bigger care process that does have evidence. But if you look at one specific component of it, you may not be able to prove that that prevents foot ulcers.

CO-CHAIR GOLDEN: And I would assume also that this measure would not replace care in an office setting. So, it would be a component of something that would go along with other activities in the clinical spectrum.

MEMBER KIRKMAN: Right. So, I think if the home health patient has a primary care physician, that they are still seeing the primary care physician would still be potentially held to the foot exam measures, for example.

CO-CHAIR ROSENZWEIG: I was

curious why patients who were not amenable to foot care education, such as patients with dementia and various other categories were not denominator exclusions.

MS. COOK: So, the measure actually specifies that the education can be provided either to the patient or to a caregiver, such as a spouse or another family member. So, if the patient is not able to receive the education directly, then presumably the home health agency would target it toward a care giver.

CO-CHAIR ROSENZWEIG: Thank you.

MEMBER TAYLOR: Is it possible to click on the link for the Cochran review? Because it has a little short paragraph that tells us what the review found about the evidence. It's pretty revealing.

CO-CHAIR GOLDEN: I think if you look in your book here, I think it should be in here somewhere. Maybe not. Is that it?

MEMBER TAYLOR: Can people read

it? Would it help if I read it?

CO-CHAIR GOLDEN: If you have got it, sure. Just summarize it.

MEMBER TAYLOR: Of the 12 RCTs included, the effect of patient education on primary --

CO-CHAIR GOLDEN: Page 28, by the way.

MEMBER TAYLOR: -- endpoints was reported in only five. Pooling of outcome data was precluded by marked mainly clinical heterogeneity. One of the RCTs showed reduced incidence of foot ulceration, relative risk 0.31 with 95 percent confidence interval 0.14 to 0.66, and amputation relative risk 0.33 with a confidence interval of .15 to .76. During one year follow-up of diabetes patients at high risk of foot ulceration after a one-hour group education session.

However, one similar study with lower risk of bias did not confirm this finding. Relative risk of amputation 0.98, 95

confidence interval of 0.41 to 2.34, relative risk of ulceration 1.00, 95 percent confidence interval of 0.70 to 1.44. Three other studies also did not demonstrate any effect of education -- likely underpowered.

Patients' foot care knowledge was improved in the short-term in five of eight RCTs in which this outcome was assessed as was patients' self-reported self-care behavior in the short-term in seven of nine RCTs. Callous, nail problems, and fungal infections improved in only one of five RCTs. Only one of the included RCTs was at low risk of bias.

CO-CHAIR GOLDEN: Any other comments or questions on this?

MEMBER SHWIDE-SLAVIN: Just as Sue noted, one of the problems is that we are talking about home health care and these studies are not in home health care.

I just found another supporting article when I was trying to find out more about foot care education and there is a

randomized controlled trial that was published in '89 of 203 patients that were randomized into groups that got education and didn't get education. And the group with education had a p value of 0.0005 for the significance of the education.

So, I think there is -- no, amputation. I'm sorry. So, they were looking at the incidence of lower extremity amputation in diabetic patients, exactly what we are looking at in these foot measures.

MEMBER KIRKMAN: The ulceration rate.

I think another problem with all these studies is it is a little bit like education about kidney disease and trying to show that you lower the rates of dialysis. You know again, it is part of a comprehensive thing. Yes, I agree that the evidence is fairly mixed, if you try to pull out patient education alone.

CO-CHAIR GOLDEN: Are we ready to

vote? Bill, you have one last comment?

MEMBER TAYLOR: Yes, the patient education that is studied in the studies are things like a one-hour group. And we are talking about somebody at a home health agency checking off a box that they did, this required foot education.

MS. COOK: Can I just clarify? The documentation on the OASIS instrument is saying that there is evidence in the clinical record that the intervention was performed. The intervention itself in many cases is significantly more extensive and it can even include, over time, a nurse might be visiting the patient every week for multiple times a week and can include both an initial information session and then also follow-up with that patient.

Again, CMS doesn't dictate exactly the education that the caregivers must provide but at least some of the home health agencies do provide much more significant education

than just a one-hour group session.

CO-CHAIR GOLDEN: Ready to vote? You have the last question.

MEMBER HAYDON-GREATTING: I'm sorry. If they don't dictate what type of education, does CMS expect a standard, like a nationally recognized program like what is in LEAP or what other diabetes educators that are accredited and certified are capable of doing?

MS. COOK: Can I actually ask if Deborah Dietz managed to join the fun?

MS. DIETZ: Yes, I am here. Thank you.

There is no specified national standard. However, the way that this is worded is that the agency must collaborate with the physician to come up with orders for specific interventions and then they must implement. There must be evidence in the record that they implemented those interventions.

So, they are required to discuss

with the physician what education and monitoring are appropriate for that patient and then implement that.

MEMBER SHWIDE-SLAVIN: I just wanted to point out even though when we initially looked at this and the evidence looked low, I really liked the way Bill proceeded through the algorithm. Because I think if it is rated as low, they won't continue with the measure and then home health agencies won't be accountable for doing the foot care and education.

Whereas, if we rated it as insufficient evidence with exception, then I think the measure, we can move forward in our grading of the measure.

CO-CHAIR GOLDEN: That would be correct.

MEMBER KIRKMAN: Although, again, I think we have to rate the evidence the way we think the evidence really is. Right?

CO-CHAIR GOLDEN: Sure. It is

time to vote.

MS. BAL: Voting is open.

(Pause.)

MS. BAL: We have moderate one; low four; intermediate one; and then insufficient evidence with exception 13. So, it goes forward.

CO-CHAIR GOLDEN: So, we will continue. That means that would pass for continued discussion.

So, we go to the next item, which would be performance gaps.

MEMBER KIRKMAN: So, in terms of performance gaps, they have a lot of data and the average performance on the measure is quite high. It is 93.4 -- I was reading that. That's okay, I can read it from here. The average performance on the measure is 93.4 percent. There is a basically a 17 percent gap between the 90th and the 10th percentile. And a 7.7 percent gap between the 25th and 75th percentile.

They do present results from 2010 to 2013 and there has been I think it is a pretty slight improvement but some improvements. And all groups in the disparity stratification are above 90 percent. And for 2013, the 25th percentile was above 90 percent.

So, I don't think there is a large gap because basically the agencies are all doing quite well in this measure. I mean again, I think it is sort of set up to be -- it is almost mandatory that the box be checked. You know I am sure there is a lot of documentation behind that but I think that this is one where it is fairly straightforward to meet the measure, technically.

So, it is hard to say from these data that there is a large gap in care or opportunity for improvement.

CO-CHAIR GOLDEN: Question for the developer, CMS. This is embedded into OASIS. Is that correct?

MS. COOK: That's right.

CO-CHAIR GOLDEN: So, regardless of, I mean if -- there are elements in OASIS that are not NQF measures. Correct?

MS. COOK: Exactly, yes.

CO-CHAIR GOLDEN: So, if there isn't a performance gap, it would be up to you whether you wanted to continue measuring it or not really. So the impact of the NQF measure per se would be for, I guess, maybe for incentives or things?

MS. COOK: So, CMS prefers to publicly report measures that are NQF-endorsed. So for the most part, OASIS-based measures that are not NQF endorsed are only reported confidentially to the agencies. And we have certainly seen that with these measure of various types that look at interventions, the improvement, since the adoption of OASIS seen in 2010 has been much more substantial for those that are publicly reported. So, the agencies do seem to be motivated by public

reporting.

CO-CHAIR GOLDEN: Other comments or questions on performance gap? Down at the end, yes, Patricia.

MEMBER McDERMOTT: Based on what was just stated, that there is an order first from the doctor and really it is based on what the doctor's order that the home health agency is doing what the doctor ordered around education and foot exam, I think that is what I heard. This is really their documentation that they followed a doctor's order. Did I misunderstand?

MS. DIETZ: The only thing I would add is that they, in home health, many of the physician orders come at the behest of the agency. So, that it means that they have pursued this line of inquiry with the physician, which got an ordered and then gone ahead and implemented.

CO-CHAIR GOLDEN: I believe that your issue about how -- and denominator

inclusion, that would be a specification question about how the denominator gets created. Is that your question?

MEMBER McDERMOTT: Well really if it is an exercise in knowing that you followed what a doctor asked you to do, it is not something they are initiating on their own. Right?

MEMBER KIRKMAN: But they have to get the doctor's order, too. In other words, they can't just assume that the doctor is going to order it. So, they have to actually sort of seek out the order, is my understanding as part of to meet the measure.

MEMBER SULLIVAN: At least I know for our enrollees, we ask the doctor for the orders. We tell the doctor what we want in the orders. Now, the doctor is completely free to disagree with us but if you leave it to the doctors to make the order, it is likely that nothing would be ordered. I mean, it is just not what they are thinking about, so that

is just now how it works.

I mean my feeling, I just want to say just about the doctors' orders but also about doing this, if it went -- I mean even though there is no gap and that all the agencies are doing it, if it wasn't in the OASIS form, they wouldn't be doing it.

I know they could but I am just saying if it wasn't in the form, it wouldn't happen at all.

CO-CHAIR GOLDEN: It helps direct activity. Yes, Bill.

MEMBER TAYLOR: I'd like to clarify what that means about being in the doctor's order. I am a primary care doctor and I only take care of about 400 patients. Most primary care doctors take care of between 2,000 and 3,000.

I send a patient home from the hospital or somehow they get a referral to home health agency. What happens as a consequence of that is I get a form in the

mail sometime later, telling me to certify the care for a certain period of time.

I have only said the patient needs to go home and have services and I get about a seven- to ten-page form with all sorts of boilerplate written in it and I have to sign my name and write a date at the bottom of each one of those pages. I actually physically see each page because I have to turn it. Do I read it? Do I know if the home health agency has included diabetic foot care in there, which I am sure they have because we now have 94 percent compliance? I have no idea if that is what I am signing.

CO-CHAIR GOLDEN: So let me just, before we get too deep into this because I think we are mixing up a little bit specifications and performance gap. So, I guess the question is, does this measure include anybody with diabetes and home health care or anybody with diabetes in home health care with a doctor's order?

MS. COOK: It is anybody in home health care with diabetes, with the very rare exception of those who have bilateral amputation.

CO-CHAIR GOLDEN: Okay. So, the doctor's order is somewhat irrelevant because if the patient --

MEMBER KIRKMAN: Well, they all have doctor's orders typically.

CO-CHAIR GOLDEN: I understand.

MEMBER KIRKMAN: Or, if the patient dies while under home health --

MS. COOK: Yes, that is quite rare, too.

MEMBER KIRKMAN: Right.

MEMBER TAYLOR: In only mention that because it was a question if this communicates to the doctor and he makes a plan of care and so on, I wanted a little reality of how, at least in one person's practice that works.

CO-CHAIR GOLDEN: Are we ready to

vote about performance gap? Okay.

MS. BAL: Voting is open.

(Pause.)

MS. BAL: We have moderate 11; low eight.

CO-CHAIR GOLDEN: Okay, that puts us in the 58th percentile. So that is continued. Okay? Not a ringing endorsement but continued.

So, high priority.

MEMBER KIRKMAN: So the group agreed that the condition is high priority, you know amputations, diabetes is high priority. You know, I think it is a little bit indirect thinking but I think presumably the home health patient with diabetes is a very high-risk patient. So, I think potentially it could be a high priority measure. But it is hard to say.

So, some one of the comments was the condition is relatively common and severely impacts quality of life, cost, and

life expectancy. However, I am not sure the measure is evidence-based, nor that performance gap remains justified.

And somebody else said it addresses a health concern that causes significant morbidity, contributes to mortality and adds to cost.

CO-CHAIR GOLDEN: Comments or questions about priority. Ready to vote?

MS. BAL: Voting is open.

(Pause.)

MS. BAL: Okay, we have high nine; moderate eight; low one; insignificant one.

CO-CHAIR GOLDEN: Okay, liability.

MEMBER KIRKMAN: So, they did present reliability data. They have a very high the beta binomial method value of 0.7 or above is considered acceptable. And theirs was 0.92 is the mean. And they were able to look at interclass correlation coefficient for agencies that have at least 40 valid episodes of care and felt that most of the total

variation is between agency variation, which would be what you would want to see. But if you are comparing agency to agency, that is where most of the variation is.

So, I think we felt like it was a highly reliable measure.

CO-CHAIR GOLDEN: Ready to vote?

MS. BAL: Voting is open.

(Pause.)

MS. BAL: High 17; moderate one; low one.

CO-CHAIR GOLDEN: Now we get to validity.

MEMBER KIRKMAN: So the developer did empiric validity testing comparing performance on this measure with other publicly reported measures for home health agencies and found some significant correlations between performance on this measure and other measures like a pressure ulcer plan and a slight negative correlation with emergency department visits. And then

they also have some face validity reports that they felt were high.

So, is this where we talk about exclusions? So, I got a little confused on the phone call about whether CMS is removing an exclusion that was previously there. But I guess before there was an exclusion, previously there was an exclusion for long-term care episodes. And I think it was thought by NQF that it would be more difficult to collect if someone was in home health for two years. But I think the developer says that it is collected once at the end of the episode, so it doesn't really matter the length of the episode. So, they didn't really think the exclusion was necessary. Correct me, if I am wrong.

MS. COOK: Right. And we are proposing to remove that restriction. So, the measure as submitted to the committee would include all home health episodes, regardless of length.

CO-CHAIR GOLDEN: To be clear, the measure as submitted includes that now but at some point you would be changing that?

MEMBER KIRKMAN: No, it is not in there.

MS. COOK: No, the measure as submitted does not include that. Once we finish this process, we will instruct the measure implementer to remove that exclusion and their calculations.

MEMBER KIRKMAN: So, the currently endorsed measure has the exclusion but you are proposing -- this measure that we are voting on now does not have the exclusion. Too many negatives here, but yes.

MS. COOK: I don't have the time line in front of me but I believe the publicly reported measure on Home Health Compare right at this instant has the exclusion but as early as the end of April, which is the next Home Health Compare release, that exclusion can be removed.

MEMBER KIRKMAN: So anyway, I think the committee felt the validity was high, reasonably high.

CO-CHAIR GOLDEN: Vote.

MS. BAL: Voting is open. Sorry, hold on. Too many -- now open.

(Pause.)

MS. BAL: Okay, we have high 15; moderate four.

MEMBER KIRKMAN: Feasibility is next. So, this seems to be a highly feasible measure. Again, it is a very easy measure because the ultimate collection is a checkbox on the OASIS form, which they are required to submit anyway. So, essentially, they have no missing data and it has been collected with the greatest of ease.

CO-CHAIR GOLDEN: I see people voting already but we are not open yet.

(Laughter.)

CO-CHAIR GOLDEN: Another false start and you will lose your medal. I think

we are ready to vote.

MS. BAL: Voting is open.

(Pause.)

MS. BAL: All right, high 19.

MEMBER KIRKMAN: In terms of usability and use, it has already been in use for three or so years. It is publicly reported. So, it is felt to be highly usable. It is also used for internal quality improvement. There was some comment in the workgroup that a potential unintended consequence would be the time and attention spent on patient education and foot care might be better spent on something else. That was someone's comment.

CO-CHAIR GOLDEN: Comments, questions? Vote time.

MS. BAL: Okay, voting is open.

(Pause.)

MS. BAL: We have a couple of people missing. Could we all just click one more time? Thank you.

(Pause.)

MS. BAL: All right, we have high 12; moderate seven.

CO-CHAIR GOLDEN: Global endorsements.

MS. BAL: Voting is open.

(Pause.)

MS. BAL: The final result is yes, 17; no, two.

CO-CHAIR GOLDEN: Thank you, Susan.

I was going to turn this podium to you, Jim, but you are doing the next discussion.

MEMBER DUDL: Is that the hyperglycemia?

CO-CHAIR GOLDEN: It is the hyperglycemia, 2362.

MS. TIGHE: Operator, do we have the developers for Measure 2362 and 2363 on the line?

OPERATOR: The lines are open.

MR. CAMPBELL: Yes, this is Kyle Campbell.

MS. TIGHE: Okay, great. Thank you.

MR. CAMPBELL: Can we go ahead with the opening?

CO-CHAIR GOLDEN: Yes, please.

MR. CAMPBELL: Okay. Well, good morning. My name is Kyle Campbell and I am a pharmacist and Executive Director for the CMS Medication Measures Special Innovation Project at FMQAI.

Our project is tasked with both maintaining and developing medication-related measures for CMS and specifically for the hospital study, we are tasked with developing new electronic health record-based measures that address adverse drug events and adverse events in medical care.

The measures have been in development for the past two years and have undergone a very rigorous development and

testing process, guided by a technical expert panel as specified in the CMS measure management system blueprint.

The measures were developed in partnership with a large academic medical center and tested in eight hospitals across the country that were selected for key criteria, such as the size, teaching status and the EHR vendor which they use. We found the measures to be feasible across all the hospitals in which they were tested.

These paired glycemic control measures that are submitted for your consideration today are the first eMeasues to be completed under a project. And in terms of importance, inpatient glycemic control has been identified as a major gap in the NQF portfolio for both inpatient and outpatient hyper and hypoglycemia have been associated with poor outcomes in the hospital setting, including increased morbidity and mortality.

Regarding the evidence per

direction from the workgroup, we have submitted a briefing document with additional studies cited to supplement our submission to demonstrate a strong relationship with glycemic control and in-hospital mortality.

Furthermore, to underscore the importance, the measures have been recommended in the National Action Plan for ADE Prevention recently published by the Department of Health and Human Services. The specifications are precise and the measures are paired to avoid unintended consequences associated with reporting either measure alone.

For hyperglycemia, to capture the at-risk population, we identified patients with diabetes through diagnosis or drug proxy and were inclusive of patients with a blood glucose greater than 200 milligrams per deciliter. The measure focuses on sustained hyperglycemia of greater than 200 milligrams per deciliter defined as two measures at least six hours apart, rather than a single incident

to reflect a validated mechanism of hyperglycemia on morbidity and mortality.

The threshold of 200 milligrams per deciliter was discussed extensively by the TEP and was selected since 200 milligrams per deciliter was generally the highest threshold considered by clinical practice guidelines and primarily the upper limit that has been used for safety in clinical trials.

Issues related to differences of testing frequency have been addressed in the measure algorithm to assure hospitals with limited blood glucose monitoring would not be incentivized by the measure.

And finally for reporting purposes, our TEP suggested stratifying the measure by ACU, ICU, med versus surge, and patients receiving high-dose steroids.

For the hypoglycemia measure, the measure identifies for public reporting severe hypoglycemic events, those less than or equal to 40 milligrams per deciliter. And this

threshold is consistent with the definition of safety outcomes in the majority of clinical trials that have evaluated the effects of glucose controls on morbidity and mortality. Severe hypoglycemic events are potentially life-threatening adverse drug events that can largely be prevented by appropriate monitoring and glycemic control.

I want to thank you for your consideration of these measures and we look forward to your review.

CO-CHAIR ROSENZWEIG: Okay, thank you. This particular measure is an inpatient intermediate outcome measure. And it basically identifies patients in the hospital. And if within a given 24 or a given daily care you have two or more blood glucoses that are measured within a four hour period was it -- I'm sorry -- six hour period, then that particular day is considered to have hyperglycemia or was counted as being one of the percentage of days, one of the particular

parts of the days that is considered to have a hyperglycemic episode.

Now, there is no question that -- I mean they present a lot of evidence here and there is quite a bit of evidence showing that hyperglycemia within the hospital setting is associated with a number of adverse outcomes, including morbidity and mortality. But a lot of the -- most of the studies that look at this define hyperglycemia in different ways. So, none of them specifically define them exactly in this particular manner, the blood glucose over 200 twice over a six-hour period.

So, the question is does the evidence for hyperglycemia being a negative outcome or negative intermediate outcome is quite clear, I think from a lot of the data from the analyses. The issue you would have to decide is whether this particular method of measuring hyperglycemia during the course of a hospital stay is an appropriate one and an adequate one, basically as an intermediate

measure for identifying greater versus less than greater hyperglycemia.

They specifically exclude patients who were admitted with diabetic ketoacidosis and hyperglycemic nonketotic state, which are obviously situations in which the patient comes in with fairly significant hyperglycemia and is supposed to be treated for this. However, I think that it may be that it is reasonable to exclude these because it is apples versus oranges. These are people who, by the very nature of the situation are being admitted for uncontrolled hyperglycemia, whereas, in the other situation you are looking for patients who come into the hospital for other reason and develop hyperglycemia.

But still, patients with hyperglycemia I think, to a certain extent, patients with hyperglycemia nonketotic state or a DKA would benefit from improved control. But there is limit to a certain extent as to

how fast you can bring the blood glucoses down and under control. And that may be more related to severity of illness than it is related to the degree to which the patient, the optimal degree to which the patient is cared for.

So it seems like those are, to me, reasonable denominator exclusions.

So with respect to evidence, I think the evidence for control of hyperglycemia in a number of different settings is very, very high. The question is whether there is evidence for the use of this particular way of identifying hyperglycemia, that is the percentage of days in which you have two blood glucoses greater than 200 in a particular six-hour period, whether or not that is the best way of being able to identify patients with hyperglycemia. And I would say that would be at a lower level of evidence.

Yes?

MEMBER CURRY: I'm having

difficulty understanding the measure, as it is calculated. So, in the numerator we are talking about a sum of percentage of hospital days. In the denominator, we are talking about the total number of admissions. And those aren't the same. In my mind, they aren't the same values. So, I have difficulty understanding how the measure is calculated.

MR. CAMPBELL: This is Kyle Campbell with FMQAI. I am going to defer that question to Almut Winterstein from the University of Florida.

DR. WINTERSTEIN: I didn't hear the last part, Kyle. I just had to unmute myself but I think you deferred to me. Right?

MR. CAMPBELL: Yes, that is correct. And did you hear the question from the Steering Committee?

DR. WINTERSTEIN: Yes. Yes, I did hear the question.

MR. CAMPBELL: Okay.

DR. WINTERSTEIN: Yes, this is one

of those measures and there are several measures in the NQF portfolio that are constructed in a similar fashion. And it is a little bit difficult to think through this at the very beginning.

Essentially when you look at the numerator and denominator, that gives you the average of the percent of days for each patient that was hyperglycemic. So, essentially the measure is constructed in two stages. The first stage, we look at every single patient's admissions, take the total number of days for this patient, identify the days that are hyperglycemic, according to our definition, and that gives the percentage for an individual patient.

And then all those percentages are summed up and divided by the total number of admissions, which basically gives us the average number of hyperglycemic days across all admitted patients. Does that make sense?

CO-CHAIR ROSENZWEIG: Well, I

think Dr. Curry is raising an interesting point. Because couldn't you just basically identify the total number of admission days the person has had within a given year and divide that by the total number of days that have hyperglycemic episodes? I mean, you don't have to necessarily divide it up by individual admissions, unless I am seeing something that --

DR. WINTERSTEIN: That's right, yes. I mean basically, we could either report the percent days across all hospital days that were hyperglycemic or we express it on the level of individual patients, essentially.

We decided on the latter, from the reporting standpoint because typically when you look at information that evaluates how consumers interpret measures, it is easier to get your arms around something that is expressed on the level of patients than on the level of days.

So right now, the measure would

basically tell me as a patient if I am admitted and I am at risk for hyperglycemia because I have diabetes or because I am starting on steroids or I have post-surgical catabolic stress, then my individual risk for becoming hyperglycemic is roughly 20 percent of all my hospital days might be hyperglycemic or 30 percent.

In contrast, if we went with the database measure, we essentially, a consumer would say okay, across all patients who are hospitalized, roughly 20, 25 percent are hyperglycemic. So it is just a matter of trying to make the information more palatable.

CO-CHAIR GOLDEN: So I am going to step in here for a second just to shape the conversation, if we could. And kick me in the shins if I got this wrong. On the table here is a question of evidence related to the measurement of hyperglycemia and whether hyperglycemia is something that has impact, you know whether glucose over 200 while

hospitalized is something to measure.

The last little bit of conversation is really into how it was reported and how it is used. Because the data is collected and then it gets constructed into a measure.

So, I was going to say that we could talk about the construction later but on the table right now is the evidence about hyperglycemia itself and the measurement of it and the definition of what is hyperglycemia, as opposed to the reporting of the measure.

Yes, Tracy.

MEMBER BREEN: Thank you. So full disclosure. I oversee inpatient diabetes operations in my health system. And so we have spent a lot of time on this concept of data collection. And what we are really talking about is glucometrics. That is the wonky term, glucometrics, how you identify hyperglycemia in the hospital and how you define it.

From my understanding, there has been no really good studies looking at the best way to define glucometrics. This is one proposed method but the methodology of identifying glucometrics in the hospital is still really being worked out. And there are lots of different ways to do things.

So that is my concern. Number one, I think the data that has been presented is about the risk of hyperglycemia. But this measure is also specifically saying this is how they are going to define hyperglycemia in the hospital with their glucometric methodology. So that is my concern, number one.

My concern, number two, is there is no benchmark for where this should be. So if a measure reports something and a hospital says well we are a 30 percent hyperglycemic, according to this and another hospital says we are a 50 percent hyperglycemic, I am not sure that we have a gold standard or benchmark to

say where that should even be.

And so to put up a measure without a benchmark is my concern, number two.

I had a third concern but I am losing it. So, I think we should really spend some time on the evidence about this.

CO-CHAIR GOLDEN: Your comment is there is a question about a new word for me, glucometrics.

MEMBER BREEN: Glucometrics.

CO-CHAIR GOLDEN: And so is there data about glucometrics as something to be measured? But then you also say that there is no evidence that there is a standard of a number 200 versus 250 and also the frequency.

MEMBER BREEN: And also how you define. So, in this concept of glucometrics, just to give the group some idea of how many different ways you can do this. So, some systems say well, you dump the first 24 hours, because that is the most active part of admissions. So, you don't even start the

clock until 24 hours after admission and then you start the clock. So there are patient-specific glucometrics, which this is. It relates to the patient. But again, if a patient has a two-day stay versus a 12-day stay and on that patient's two-day stay they have one hit on one day, they come up with a 50 percent hyperglycemic rate; whereas, someone with a 12-day stay who is hyperglycemic maybe on three days had more hyperglycemia, their percentage comes out lower. So then you get into the what is a weighted glucometric scale? How do you make it patient-specific?

It has been so hard because there are so many data points. We know so in hospitals point of care testing ranges between 5,000 to 30,000 tests per year. It varies in some patients, they get so many tests, so there is so much data and noise we have all had a lot of struggle to define how to make that very patient specific.

CO-CHAIR ROSENZWEIG: Yes, I think that gets directly to the point of what Dr. Curry was mentioning as well. If a person is in the hospital for 20 days and has hyperglycemia for a certain percentage of days, let's say 40 percent out of those days have two blood glucoses greater than 200, that is counted as one admission. And another admission of two days would be counted as a separate admission. And as far as I could tell from the way this is calculated, each would be weighted the same. And it might be --

Now this doesn't directly pertain to evidence except if it does because we know that there is good evidence for a relationship of hyperglycemia -- extent of hyperglycemia to poor outcomes. The question is whether or not we are measuring that within the context of this measure.

MEMBER BREEN: And we also don't have -- we have data about hyperglycemia in

very specific sub-populations of patients, critical care, cardiothoracic surgery patients. We don't have good data on the general med surge population in inpatients for a potential single day of hyperglycemia as this would potentially define. So there is no data to say someone coming with an elective knee repair who is in the hospital for three days and has a hyperglycemia hit on one of those days so they come up with a 30 percent rate, we don't have data to say that that has some associated risk for them for a consumer to look at a hospital report and say oh, I have got a 30 percent risk of having hyperglycemia if I got me knee done.

I'm just uncomfortable.

DR. PACE: Would they be in the denominator, unless they -- I mean --

MEMBER BREEN: Yes.

DR. PACE: If they don't have diabetes?

MEMBER BREEN: But I don't know.

But if the measure is also saying this is how they are defining hyperglycemia, I think that has got to be a -- I mean we can get to it there but the data is not just about hyperglycemia. It is about how they are measuring and defining hyperglycemia in this measure.

CO-CHAIR ROSENZWEIG: Exactly. And a lot of the other studies don't define hyperglycemia by two blood glucoses greater than 200. A blood glucose greater than 200 occurs quite frequently in the setting of all sorts of different situations. A person getting an IV stick in the wrong place, a variety of other things.

The way the data would be collected in this measure couldn't distinguish between all of those things.

A blood glucose, let's say, then in addition, how would you weight a blood glucose of 350 versus 202. So, if you had two blood glucoses of 350, they are counted in the

same way as if you had two blood glucoses of 201. And perhaps there is a big difference between the two.

Yes? Oh, I'm sorry.

CO-CHAIR GOLDEN: Oh, that's okay. Let me go to the far end.

MEMBER DUDL: Well, just to ask Tracy, do you think then this fits the issue? I mean we know it is important to measure. We don't know exactly how to do it. There is not proof. Is this insufficient data but with exception?

MEMBER BREEN: I'm so conflicted about this because hospitals need to show how they are doing. Right? So, I can say from my experience in many, many hospitals, hospitals have no idea how they are doing. And there is this vast treasure trove of data that they are sitting on that nobody looks at. Right? It is where all those 50,000 data points of point of care the patients suffer through, get their finger tested. Nobody looks at them or uses

them to manage care.

So, we are struggling with that. And when you begin to look at it, then you unearth that you have huge chunks of patients who are hyperglycemic and potentially at risk but we are almost, we are so far behind that without having hospitals begin to look at the data, we don't even know where we are.

That is my concern here. We don't even have a good baseline assessment of where hospitals are right now to say that okay, we are all here, so we should then get here. And this measure begins to get at that at least in some way but again, there is no benchmark associated with it. And we are going to jump to unintended consequences but of course the unintended consequences of that are --

CO-CHAIR GOLDEN: Sue?

DR. WINTERSTEIN: This is Almut Winterstein. It is very unfortunate that we were not allowed to attend the meeting in person.

CO-CHAIR GOLDEN: Can I hold you off until -- I have a couple of committee members and then we will have you chime in.

DR. WINTERSTEIN: Okay. I was asking there were a few clarifications, I think we would provide. So yes, as long as you have us in mind, it might help the discussion.

CO-CHAIR GOLDEN: Okay.

MEMBER KIRKMAN: So, maybe I'm just being too concrete here but I still think the evidence is really the evidence about that hyperglycemia is bad and that there is some evidence that managing it is better.

I think a lot of this really is about the specifications. And I agree that that sort of relates to evidence but we know that there is not a consistent way that things are measured in all these different studies.

So I mean I just sort of feel like we need to just vote on the evidence and then move on to talk about specifications because

otherwise I think we are getting too much into other things.

CO-CHAIR GOLDEN: There is a fine line here. I mean I could see Tracy's point. But I can see otherwise we will end up --

MEMBER KIRKMAN: And then in terms of a benchmark, I mean I sort of agree with you but on the other hand, we all kind of agree that a lowered number is better. And so you are going to have all these publicly reported hospitals and everybody is going to try to move down. I am not sure that it is really necessary to say that it should be less than ten percent but maybe I am wrong.

CO-CHAIR GOLDEN: So I am going to go to Bill and then I am going to go to the phone.

MEMBER TAYLOR: Thank you. Until Sue just mentioned we know that higher glucose is bad and Jamie gave us the powerful evidence that high glucose is associated with mortality and so many other adverse outcomes but Sue

said and we know that controlling it is good.

You endocrinologists --

MEMBER KIRKMAN: I didn't say that.

MEMBER TAYLOR: You didn't. Okay. So I would love to hear a little about the evidence that we are actually helping people when the high glucoses are controlled. Because my understanding, and I am far from an expert in this area, is that there have been a lot of disappointing studies that have shown that when you control more tightly that you don't improve outcomes. And the concern is that high glucose may be a marker for being sick and having a bad outcome, rather than that the control of it is necessarily going to improve things.

And if we are using this as a marker of quality and say you are a bad hospital if you have a lot of people with hyperglycemia, are we in danger of saying you are a bad hospital because you take care of a

lot of sick people?

CO-CHAIR GOLDEN: I'm going to give this to Jamie to discuss.

CO-CHAIR ROSENZWEIG: Yes, I mean this speaks to a lot of the major studies have been in the ICU setting and it has been going on -- there has been a controversy that went on for quite a number of years, starting with the van den Bert study in Belgium in which they were able to show -- and there were some differences between surgical ICUs and ICU settings but it was clear that some studies seem to show that evidence of very tight control would be beneficial with respect to outcomes, both in the hospital and then post-hospital. But subsequent studies didn't bear those out.

And I think the question was mostly settled by the nice sugar study, which came out a couple of years which was a large multi-center study, international study that looked at patients with very tight control,

compared with patients with less tight control. But both of those population groups were still lower than blood glucoses over 200.

So the difference was between keeping blood glucoses being between 90 and 110 and the keeping them between 110 and 180. And between those two groups it looked like perhaps the patients in the 110 to 180 group actually had fewer adverse outcomes or at least mortality-related outcomes.

So, that was a question of very tight control but I think most of the data would say that patients with hyperglycemia, per se, both in the hospital -- certainly within the ICU setting, if it is defined as greater than 200, that that would constitute potentially association with adverse events.

CO-CHAIR GOLDEN: I am going to go to the phone and then I will go to the other folks. So, there was a comment from one of the developers on the phone.

MR. CAMPBELL: Yes, I would start.

This is Kyle Campbell with FMQAI.

Just in regard to the benchmarking question, I think you know hospitals, we aren't setting a specific benchmark. More than likely, this would be reporting on how hospitals are compared to the mean of all the hospitals within the country. And those that are statistically significantly above the mean and those that are statistically significantly below the mean and we think that this would start, as mentioned by the Steering Committee, there is this treasure trove of data and we do know there is a clear quality gap here. And this would start the conversation about moving toward better glucose control.

In terms of the specifications, we did wrestle a lot with how the day should be measured, what day should start, what day should end. And I would just defer again to my colleague at the University of Florida, Almut Winterstein, for a little bit of an explanation of that.

DR. WINTERSTEIN: Okay, thank you, Kyle.

Yes, I very much appreciate the concern about glucometrics and I hadn't heard that term before either but I think it is a pretty good term.

If you look at the clinical trials that have provided evidence on this, they usually average their glucose results, which makes sense because they follow a clinical trial protocol, so there was a number of tests available for a given day and, therefore, an average will probably will fairly well capture what is going on in a particular day.

And if you look at van den Berg, this was one of the studies that was quoted earlier. The average glucose, I believe, for the control group, was somewhere in the 180s, I think 178 or something like this. And then for the intervention group it was at 105 or 108.

And as was already pointed out, I

am not aware of any trial where the upper limits in the control group was below 200 -- sorry -- above 200. So, usually, the titration regimens for patients in the control group was set as such that insulin had to be introduced or had to be increased when glucose values exceeded 200.

So, this is kind of where we were coming from when we were trying to construct the measure here.

Now, the thing is, of course, in a hospital environment in real life we don't have clinical trial conditions and we don't have the protocols that prescribes how often glucose would be measured. And if they are indeed if there is just one or two or three values and we average those values, we will definitely remove the extremes from the overall glucose management. An average is always, obviously, in the middle of two values. So, we wouldn't really properly capture whether they really were hyperglycemic

events.

At the same notion, it is fairly clear from the pharmacological perspective that sustained hyperglycemia is really what the issue is and that single peak is probably not the important part.

So, this is why we decided to propose to go over the measure that would capture sustained hyperglycemia, defined as having at least two hyperglycemic events that are at least six hours apart, suggesting that there were a longer time period where glucose values at a given day exceeded even what the control group in a clinical trial setting.

I think the other part that is important to mention here is that we were trying to deal with an area that hospitals essentially didn't measure. So we have in the measure a numerator, another parameter that deals with situations where just one glucose value is available that is above 200. In those instances, we consider that also

hyperglycemic with the idea that if nobody measured again, very unlikely was there anything done about trying to get hyperglycemia down because typically when I am introducing insulin or increasing insulin dose, I would go measure again. So, the idea was not to incentivize hospitalize hospitals that essentially ignore an elevated glucose value.

CO-CHAIR GOLDEN: Thank you very much. We should soon get to vote about evidence. Evidence is going to be about -- and we are going to get back to specifications and a variety of other issues. So, we will be reprising a lot of this conversation.

I am going to go Jessie because she hasn't said anything. And then I will go to Tracy.

MEMBER SULLIVAN: I'm sorry. I'm going to ask the developers to answer Bill's question, which was because I just don't know. Is there evidence that a better control of

hyperglycemia in the hospital leads to better outcomes? Because I think what we just heard, if I understood what you said, James with that, in the nice sugar control, slightly looser control is better than very tight control, if I understood that right. But what I didn't hear is keeping the blood sugar below 200 better than -- or is high blood sugar just a marker of sick people?

CO-CHAIR ROSENZWEIG: No, I think there is a lot of evidence suggesting that keeping blood glucoses under 200 is beneficial. The issue is more related to -- the issues that came up were more related to under 200 if blood glucose is very, very tightly controlled versus somewhat more loosely controlled, which is better. So, that is the situation.

So, I think the evidence for blood glucoses in general of being over 200 not being good is there. But I did, as I indicated before, suppose blood glucoses are

over 300 or over 400, should that be weighted differently than two blood glucoses? That is a separate issue. I don't know it pertains to the evidence but it is a separate issue.

CO-CHAIR GOLDEN: That will get into specs and so forth. Tracy?

MEMBER BREEN: To make the discussion about evidence strictly related to hyperglycemia in the hospital, I think there are two issues we are talking about. One is maintaining control, which is doing something to prevent high glucose. The only interventions that have shown any benefit have been in the ICU setting. There has been no good evidence in the non-ICU setting that an intervention is going to really have good impact.

Now, there has been association data in the non-ICU setting. So, there have been associations to say the higher glucoses in the general med-surg populations are associated with increased length of stay,

catheter-associated infections, potentially poor wound healing, but there haven't been the same level of detailed studies on interventions in this non-ICU population.

So all the intervention studies that the developer is presenting, especially in terms of the data collected have all been ICU-based populations. So again, I just had a concern about using this ICU-based data and applying it more broadly, especially with the data collected in terms of risk. So, even those ICU studies, each one of those patients are having their blood sugar monitored very one to three hours. That is a lot of data points in one day, as opposed to a non-ICU population where patients might have two values, anywhere from two to four or five. And so I don't know how to compare those two.

So again, forgetting the way it is measured, if we are just talking about hyperglycemia in an inpatient setting, the association is there for morbidity in the non-

ICU population. But the data, as far as I know, there has been no good intervention study in the non-ICU hospital population. But if anyone has any data about that, I would love to hear about it.

CO-CHAIR GOLDEN: Sue.

MEMBER KIRKMAN: Well, I think the -- I just can't remember if it is RABBIT 1 or RABBIT 2 in the surgical patients that was a non-ICU population and did show better outcomes with sort of basal-bolus therapy and better glycemic control versus sliding scale and worse glycemic control. And they were pretty hard outcomes including, I think, mortality was actually reduced, although that may have been a fluke.

The other thing I wanted to say about the evidence is I mean I agree that ICU data is where most of the data lies. And I think we get into trouble when we say because really, really tight control was no better or was worse than moderate, therefore, there is

no evidence for treating hyperglycemia. I mean I don't think that is what Bill was saying. But it is sort of like saying because the ACCORD blood pressure trial didn't show increased benefit for less than 120, other than on a stroke, we shouldn't treat blood pressure anymore.

And the other thing I wanted to say about the non-ICU setting is there considerable consensus from multiple organizations, you know Society of Hospital Medicine, the ADA, ACE, et cetera, et cetera, et cetera, for this sort of 140 to 180 range.

So, that is --

MEMBER BREEN: But that is based on the ICU data. I agree but that is based on those numbers from the ICU and sort of outpatient --

MEMBER KIRKMAN: It is based on the ICU data and a look at the totality of evidence. So, there is a lot of -- and I agree, it is lower level evidence.

And this performance measure is really looking above that. I mean, I agree if it sort of said the percent that are between 140 and 180, that would be taking it a little bit too far. But I think that some of the evidence is reasonable.

CO-CHAIR GOLDEN: I think after Jamie, I would like to try to vote. Helen, do you want to comment?

DR. BURSTIN: Just one quick comment about the lack of benchmarks. Because this has come up in other measures and it is pretty common to have measures that are viewed as adverse events, for which there is no number as to what the right number is. Lower is better, higher is better, depending on what you are talking about. Adverse events it is certainly lower.

But very similar discussions we have had over the year about C-section rates, for example. So, I don't want that to feel like that is an issue that would really focus

on the evidence. I think the benchmark issue is a little bit outside the scope.

CO-CHAIR ROSENZWEIG: Just with respect to the evidence, Tracy, in Umpierrez's article 2012 in the non-critical setting, it did make recommendations for generally better control. You could say the evidence may not be strong but it was there.

MEMBER BREEN: Yes, I and I respect and have sat on many of the committees saying this is good because we know in our heart of hearts that it is good. But if we are saying is there evidence, right, so we have our expert opinion, our recommendation evidence and that is a lower, you can say whatever grade that is.

But in terms of intervention done in a non-ICU setting, a lot of these interventions have also just compared two ways of controlling hyperglycemia. It is not like they have said okay in one unit you are just going to do whatever you are going to do and

allowed to be hyperglycemic and in another unit you are going to control. They have really been looking at two different ways of controlling glucose and finding out which is the best way to control the glucose. And we are getting way into the weeds on this. I know this but I just -- yes.

CO-CHAIR GOLDEN: I am trying to get us to a vote. Bill, do you have something else?

MEMBER TAYLOR: Is there a way we are supposed to proceed if we feel like the information we have been given about the evidence is inadequate? I mean you endocrinologists know a great deal about this and I think we can all trust you.

CO-CHAIR GOLDEN: Well, I mean you have got the vote coming up would be you think the evidence is high, moderate, low, insufficient, or insufficient with exception.

MEMBER TAYLOR: Well, In think the evidence we have been shown is inadequate.

CO-CHAIR GOLDEN: Okay, then you vote that way.

MEMBER TAYLOR: No, but there is evidence that we haven't been shown. I know we have to vote today and so on but it is a problem of our process that the developers provided us information that showed high glucose is bad without going into this other important discussion about and what is the evidence that bringing it down is helpful.

CO-CHAIR GOLDEN: Over here.

MEMBER DUVA: I just have a quick comment. I mean I think that this is a great discussion. I know we have time limitations but I also just want to point out that the evidence we have been considering for each of the measures has had similar issues. And so, I feel like we just need to be aware of kind of raising the bar of our expectations of evidence or what our consistency is, as we go through the meeting.

CO-CHAIR GOLDEN: Yes, Helen? Not

Helen -- Karen. I'm sorry.

DR. PACE: Just one other comment about the intervention. I mean as Helen said, with outcomes in general, we are measuring outcomes that are important. We are looking at outcomes that are associated, in this case, this is an intermediate outcome associated with other kinds of bad things.

But we don't need to know that there is a benchmark because the whole point of this is to look at in comparison and learn from that. And also, we may not know all of the right interventions. But again, that is the value of measuring outcomes, so that things that maybe people think can't be done or there may be multiple ways to address it, that that is not what we are trying to get at.

So, obviously if you think that here is really, if this is something that can't be treated or there is unintended consequences, then we need to think about that. But I don't know that those two

questions directly relate to evidence.

CO-CHAIR GOLDEN: We are going to repeat some of this about the thresholding and the normative aspects during the other pieces.

Do you really want to say something?

MEMBER BREEN: It's very brief.

CO-CHAIR GOLDEN: Okay.

MEMBER BREEN: So, it is so brief because I know we are going to talk about this later. But at the final state if we are talking about glucoses of 200, greater than 200 is bad. So I mean that is what all the data shows. We know that there is an association of risk with glucoses greater than 200. And if we want hospitals to report that, which I think we do and we don't need a benchmark, which is great, that is very reassuring. So again, if we are looking at the evidence around hyperglycemia and not about the way that they are measuring it and we are going to talk about that, then I just

wanted to clarify my position that yes, there is risk and harm associated with glucoses greater than 200. So, that is reasonable to think about.

CO-CHAIR GOLDEN: So when the fog is rolled in, the fog is rolled out, and the fog has rolled back in.

(Laughter.)

CO-CHAIR GOLDEN: So I think people are ready to vote. Let's try to vote.

MS. BAL: Voting is open.

(Pause.)

MS. BAL: Okay, we have high five; moderate eight; low one; and insufficient evidence with exception is five. So, it will go through.

CO-CHAIR GOLDEN: So we will move on and probably re-circle. Performance gap.

CO-CHAIR ROSENZWEIG: Yes, so I think it was generally recognized that there is a fairly substantial performance gap, both in the ICU setting and in the non-ICU setting

with respect to this particular -- with respect to measurement of hyperglycemia. Actually, the existence of hyperglycemia in those circumstances.

CO-CHAIR GOLDEN: I'm going to suggest we focus our discussion on this aspect to is there a performance gap or practice variation in management of blood sugars in the hospital? We can talk about thresholds and levels elsewhere, but is there practice variation in managing glucose levels in hospitalized patients?

CO-CHAIR ROSENZWEIG: Yes, I would say yes, absolutely.

CO-CHAIR GOLDEN: I just wanted to get that started. All right, ready to vote? Let's vote.

MS. BAL: Voting is open.

(Pause.)

MS. BAL: Okay, we have high 16; moderate three.

CO-CHAIR GOLDEN: All right,

impact.

CO-CHAIR ROSENZWEIG: Here again, I think we felt that the impact of -- that this actually does actually address or maybe not this measure but the issue of identifying hyperglycemia and treating it in the hospital setting. It does address a specific national health goal and priority. It is a major health goal and priority.

And the issues of the caveats of how the measure is defined and so forth in dealing with other sections but I would say yes.

CO-CHAIR GOLDEN: Comments or questions? Ready to vote? Please vote. Wait a minute. It's not ready yet.

MS. TIGHE: It is open now.

(Pause.)

MS. BAL: Sixteen high; two moderate; and one insufficient.

CO-CHAIR GOLDEN: Okay. So, now we get into -- coming up now to reliability

and validity. And I believe that -- let's make sure before we get too deep into this -- reliability would be the accuracy of what gets recorded; whereas, validity gets into some of the thresholding and so forth.

CO-CHAIR ROSENZWEIG: Yes, with respect to the reliability of collecting data of this sort, I think there are obviously problems. Some hospital systems have ability to collect all of the finger stick data and integrate them into the electronic medical record. Others are less able to do so.

And this can become a major issue. So, if you are actually not able to sample, to actually collect all of the samples, then your ability to actually identify the number of hyperglycemic episodes will be diminished. And there is also --

CO-CHAIR GOLDEN: So just to clarify, this is only an eMeasure. Correct? This is not an abstracted measure

CO-CHAIR ROSENZWEIG: Yes, it is

an eMeasure but many hospital electronic medical records don't necessarily easily -- they probably will in the future but don't necessarily easily incorporate finger stick blood glucoses completely. You would have to collect the finger stick blood glucoses from a separate database and so forth. And then how to integrate them with the lab glucoses which are also collected on different samples, blood samples.

I would just say that there may be some reliability issues. And then the other issue is the sample variation. So, if a patient is not -- obviously, if blood glucoses are not measured within a six-hour period, then you are going to be underestimating the degree of hyperglycemia.

CO-CHAIR GOLDEN: Let's go down to the far end there.

MEMBER McCOLLISTER-SLIPP: Well, I guess my question was why would that be difficult? I can't imagine. I mean my

company deals with aggregated EHR data and every data set we have looked at has static glucose measures. Now I don't know, it doesn't say whether that was taken at beside or whether that was some sort of lab value that was just a static measure.

But I just can't imagine. I mean maybe I am just naive and I have never had to work with an EHR directly but I just can't imagine that that would be difficult.

I guess one of my questions, though, as somebody who deals with this on a consumer level and is involved in these issues as a patient with FDA and stuff is the reliability of the point of care meters. I mean, is that something we need to take into consideration. I know you guys are in the process of narrowing the margin of error but if it is 202 versus 199 from a clinical perspective, that is not going to make any difference. But from a measure perspective, it would.

So, I mean it is part question part philosophical point.

MEMBER BREEN: It is about the way hospitals collect this data. So all hospitals have access to this. If you check a point of care glucose, you have access. It sits in your point of care database and there is a very hardworking, dedicated point of care supervisor who usually sits in the basement with no windows sitting on all this data.

From how that data gets to the hospitals, the lab systems are basically two ways. One, it is either hand transcribed by staff, so staff check on their ACCU-CHEK. That ACCU-CHEK data is delivered electronically to the ACCU-CHEK database, which is the RALS system. And then a medical assistant or someone else hand transcribes into the medical record. That has some errors built into it and we already know that there are transcription errors.

The second way is that there is

various patches available for hospitals to purchase at a certain cost which allows that data that is sitting in a point of care database to be transmitted electronically.

So all hospitals have access to this data. Whether or not they are using it or using it fully is another issue. But I think the data is fairly reliable, once you get to that point of care database.

The concern, potentially, that you are bring up about the reliability of the ACCU-CHEK or just the meters in general, they are typically ACCU-CHEK, but any glucose meter that is being looked at by the FDA, there is a built-in variability that is allowed for, whether it should be or not on the hospital. And that is just standard of care. So, I don't know that we necessarily need to address the meter reliability right now on this. Because that is just what we are using for clinical practice at this moment.

The question comes up, though, in

terms of the measure, in practice many clinicians at the bedside, when they get an abnormal value will repeat it within five minutes, just to make sure that it is really abnormal. And then they will repeat it again five minutes later to really make sure it is really abnormal.

And so again, when you get into reliability, if you are saying you are getting two values within six hours but two of those values were within five minutes, it is really the same value. Right? And so what does that mean in terms of how reliable it is?

CO-CHAIR GOLDEN: The bench was two values six-hours apart.

MEMBER BREEN: No, within six hours. It is within a six-hour window.

CO-CHAIR GOLDEN: Oh.

MEMBER BREEN: Maybe the developer can comment on that but it is my understanding it is within a six-hour window.

DR. WINTERSTEIN: It is six hours

apart. It is not within.

MEMBER BREEN: Six hours apart.

DR. WINTERSTEIN: Yes.

MEMBER BREEN: So, but then can you describe the window? So, on a particular day, so if we are talking about a 24-hour cycle, you are saying that two values six-hours apart in that cycle would count.

DR. WINTERSTEIN: Correct.

MEMBER BREEN: At least six hours apart.

DR. WINTERSTEIN: Yes.

MEMBER BREEN: Okay, thank you. Never mind.

CO-CHAIR GOLDEN: Sue.

MEMBER KIRKMAN: Yes, I just wanted to address Anna's comment about the meter reliability, which I think is a good point but the variation is on both sides of the numbers. So, you are going to have equally as many that are below 200, when they really were above 200. And it also is not

going to vary between hospitals.

So, if you are comparing hospital X to hospital Y, it is not like hospital X was more effective by meter variability than the other hospital. So, I think it will all wash out but it is a good point.

CO-CHAIR GOLDEN: Jessie?

MEMBER SULLIVAN: Yes, I just have clarification for the developer from the last thing that you said. If it is six hours apart, if one blood glucose is drawn on Tuesday and it is 250 and it is not drawn again until Thursday, is that two days of hyperglycemia?

DR. WINTERSTEIN: No, it is always calculated for a particular day. So, the measure requires that there are two measures at a given day there were at least six hours apart and that were 200.

The only exception to this if there were just one measure available at a given day, just one, and that single one was

above 200, that would be counted as a hyperglycemic day as well, with the idea that people essentially ignored it and didn't remeasure. And we didn't want to incentivize hospitals that don't remeasure.

MEMBER SULLIVAN: Okay, so just to make sure I understand. So, if we get 250 on Tuesday and 300 on Thursday, that is two hyperglycemic days, Tuesday and Thursday. And there was on measure on Wednesday, so that is not a hyperglycemic day. Is that --

DR. WINTERSTEIN: No, that is not correct. So if Tuesday we had a 250 and all the other values that were done on Tuesday were below 200, Tuesday would not be a hyperglycemic day because we don't have two independent measures more than six hours apart on Tuesday.

Wednesday, if there were no measures done whatsoever that would count as a hyperglycemic day, unless the patient was normal glycemic for two days, at that point,

we allow that no measurement is done whatsoever.

And then on Thursday, the same thing. If there is just one single value above 300 and all the other values are below 200, Thursday would not be counted as a hyperglycemic day.

MEMBER SULLIVAN: So I guess I had said in my example a patient is in the hospital Tuesday, Wednesday, and Thursday. During that time, they have two blood glucoses drawn, period. No other blood glucoses. One is on Tuesday and is 250, one is on Thursday and is 300. How is that counted?

DR. WINTERSTEIN: They all would be hyperglycemic. Sorry, I misunderstood you. Yes, if there was only one value that was elevated on Tuesday, that would be counted as hyperglycemic. If Wednesday there was no measurement done altogether and the previous day was hyperglycemic, then the idea is nobody looks but very likely, things have not

improved. So Wednesday would be counted as hyperglycemic and Thursday as well.

MEMBER SULLIVAN: Thanks.

CO-CHAIR GOLDEN: So again, for the developer, a quick question. If someone gets admitted at night, what happens there at night? If someone gets admitted at 8:00 or 9:00 at night and has a blood sugar of 250, are those in the calculations or out of the calculations?

MEMBER BREEN: Are they calendar days or are they 24 hours from the time of admission?

DR. WINTERSTEIN: Yes, they are calendar days. So, if somebody gets admitted after noon of a particular day, the following day is not looked at all. So, it actually is more than a 24 hour period that we would not look. We use calendar days just for simplicity to put the measurement algorithm in place.

So, if a patient is admitted

before noon at a given day, then we would start to look the following day for hyperglycemia, unless the patient came in with a blood glucose value of lower than 400, at which we would give another day to allow slower titration.

CO-CHAIR GOLDEN: So the first day is not counted. Okay.

DR. WINTERSTEIN: The following day is the block for measurement.

And then the overall follow-up that might be important to know as well, we truncate follow-up at day ten of hospital admission. So, there is no patient who would contribute 20, or 30, or 90 days to the measure.

CO-CHAIR GOLDEN: We are starting, again, to bleed into reliability and validity. But we are trying to understand what we are looking at.

Go ahead, Tracy.

MEMBER BREEN: I have a question

for the developer. So, when you define admission, are you saying admission to the hospital or presentation to the ED? Because many patients stay in the ED upwards of 22 hours. So, does the clock start ticking when they present to the ED if they are subsequently admitted or is it their official admission date?

DR. WINTERSTEIN: The admission time is set at when the admission order is written and that is consistent with the CMS definition of admission and all the other measures that look at admission time that are currently endorsed use that time.

CO-CHAIR GOLDEN: Okay, ready to vote?

MS. BAL: Voting is open.

(Pause.)

MS. BAL: Okay, we have high one; moderate 17; low one. So, we move on.

CO-CHAIR GOLDEN: So, now we go to validity. Have fun.

(Laughter.)

CO-CHAIR ROSENZWEIG: My statement here would be that in ideal situations in hospitals that have the appropriate situation validity can be achieved but that probably in a large number of hospitals currently will have difficulty implementing this and will have to put into place more extensive ways of being able to identify this.

CO-CHAIR GOLDEN: Yes, I think that is in usability and feasibility.

CO-CHAIR ROSENZWEIG: Oh.

CO-CHAIR GOLDEN: So, this is where we get into the specs.

CO-CHAIR ROSENZWEIG: Oh, I'm sorry. Okay, yes. So, I think there was a certain amount of I think -- I guess face validity is probably moderate to high.

CO-CHAIR GOLDEN: So let's make sure we understand when we have the validity. So, what we have understood now is that it is six hours apart; that the first day is not

included; that the ER values are not included.

And again the developer, maybe you can help us here, I think I heard you say that if somebody is admitted with a high blood sugar there is also a compensation for that. So, if somebody gets admitted with a blood sugar of 300 or 400 because they have been at home with that for a month or two, what happens to those patients?

DR. WINTERSTEIN: Yes, if somebody is admitted with a sugar over 400, we give another day to bring the sugars down. So typically, if somebody is admitted before noon, we would start to look the following day. If somebody is admitted afternoon, we give already another day. So, we are basically starting to look at hospital day three.

If somebody is coming in with an elevated sugar, so the first admission sugar is over 400, in that instance, we add another day in which we are not looking.

CO-CHAIR GOLDEN: And my last question is and this measure is reported -- when it gets reported, I wasn't clear just how exactly is it reported. So it is percentage of days hyperglycemic or number of days hyperglycemic?

DR. WINTERSTEIN: It is the average percent of hyperglycemic day per patient. That is how it is reported.

CO-CHAIR GOLDEN: An average percent of hyperglycemic days per -- so, if somebody had one day and they were in the hospital for -- well, --

So, if the patient is in the hospital for say three days and the first day doesn't count, you then have two days. So, you would end up with, if you were hyperglycemic on a second day, it would be one out of two. So, it would be 50 percent?

DR. WINTERSTEIN: That's right. Yes, it would be 50 percent.

CO-CHAIR GOLDEN: Fifty percent,

okay.

DR. WINTERSTEIN: But keep in mind, that is average to cross roughly 2,000 patients. So, it actually comes out very similar to averaging across all days because there are so many of those.

CO-CHAIR GOLDEN: And in your testing of the activities, what kind of average number were you seeing in terms of -- you say this is a normative measure. What is your -- what has been the experience in terms of the typical number that you get or range?

DR. WINTERSTEIN: Yes, it is roughly between 20 and 30 percent of all days. There are a few outliers that are below 20 and above 30. But if I had to characterize the interquartile range, that is roughly where it is.

CO-CHAIR GOLDEN: Comments and questions otherwise? Ann.

MEMBER KEARNS: Yes, I had a question about why it says in the denominator

that after you exclude the first day or the first two days, depending, then it is truncated at ten days.

So anybody who is in the hospital for 11 days, the 11th day doesn't count or the 12th day?

DR. WINTERSTEIN: Yes.

MEMBER KEARNS: And why is that?

DR. WINTERSTEIN: Just to avoid that an individual patient contributes a lot of time and starts to skew the measure. That concern was mentioned earlier. I mean you know we could theoretically include patients who were in the hospital for 90 days. And if for some reason one had decided that the current regimen is fine and every single day is hyperglycemic, that patient would dominate the measure to a certain extent. But just in an effort to normalize the data.

CO-CHAIR GOLDEN: Tracy?

MEMBER BREEN: I have a question for the developer about your definition of

diabetes. How are you defining or indicating patients have diabetes in the hospital? It says with a diagnosis. But primary diagnosis, secondary diagnosis, any other way you are tagging people as having diabetes in this?

DR. WINTERSTEIN: Yes, the denominator uses three ways of identifying diabetes and there is different reasons for this.

The first one is just a diagnosis of 250 at any diagnosis field. So that could be the principle or any secondary diagnoses. The second is the presence of any antihyperglycemic medication. And the third is a single blood glucose value of 200.

And I tell you why that third we included is because that might not be intuitive. That is related to those patients who develop hyperglycemia, even though they don't have diabetes in the hospital. So, that is the classic patient who is on hydrosteroids or post-surgical. Those patients, if they

were not started on anti-diabetic regimen would not be captured in the denominator in any other way.

So, a single value will flag them as being at risk for sustained hyperglycemia. And that would make it in the denominator as well.

And quite interestingly, if you take the three definitions, they are really complementary. So, there is different patient populations. Of course, there is some overlap but there is actually really different patient populations that are captured.

CO-CHAIR GOLDEN: Tracy.

MEMBER BREEN: Hi, this is Tracy. For the developer. So, in terms of you including in the definition of people with diabetes a single patient with a glucose of 200, let me imagine this case scenario.

A patient with COPD being treated for steroids has a single episode of glucose of 210. So they are now flagged as part of

your cohort. The steroid step, they get no more further glucose testing for the remainder of their three days but they are in your pool. So you are going to have three days where no glucose testing was done. The way you described it, each one of those days will count against the hospital because they have not checked a glucose and they will be assumed to be hyperglycemic.

DR. WINTERSTEIN: Yes, it would count against in hospital if there were not two days -- with the previous day was not normal glycemic. So the idea is if I had a patient who had a 210 yesterday and I don't check today, we will count that as a hyperglycemic day. That is correct.

But then again, if I had a patient who had a 201 yesterday, I will check today to see whether the glucose came down.

CO-CHAIR ROSENZWEIG: And these would be people who don't necessarily have diabetes.

DR. WINTERSTEIN: Yes. In fact, in our formative testing hospital, where we developed all of this, this is a very large tertiary care hospital with a large transplant population, large surgery population, we had a good third of patients who were not diabetic but have sustained hyperglycemia.

CO-CHAIR ROSENZWEIG: Now, in our discussion over the phone, I think you were on that call, there was the issue that was raised about mixing apples and oranges such as if a person is in the MICU for a certain number of days and then is on the floor for a certain number of days and yet all of the data is aggregated into one pool for one admission.

You mentioned something about actually collecting separate data for MICU -- for ICU versus non-ICU but I didn't see it in the measure worksheet. Could you just describe what the plans are for that?

DR. WINTERSTEIN: Kyle, would you like to respond to this or should I continue?

MR. CAMPBELL: Sure. Sure. So, our plans were to, in terms of the reporting of the measure, to stratify the reporting so that there would be a separate score for ICU versus non-ICU and med versus surg patient populations, to provide a little bit of additional information for the hospitals about the differences in those subgroups.

CO-CHAIR GOLDEN: Sue.

MEMBER KIRKMAN: So, just to go back to the single blood glucose over 200. I mean you know this better than I do but the people that have hospital hyperglycemia without a diagnosis of diabetes are actually the highest risk patients and have been included in all those studies that have generated the high or low evidence. So that is one thing.

And the other thing is that I do think -- I don't think you should have a blood sugar over 200 and then just ignore it and never check it again. So, I think this

incentivizes people to keep checking, which means if it really did go away, they would look good on the measure or to treat the patient appropriately if it doesn't go away.

So, I think that is reasonable. I mean I think, if anything, it is going to incentivize follow-up and potentially treatment for those patients.

And I think if you excluded that, then you could have the patient who has a blood glucose of 280 and everybody says oh, well it is just the steroids. We are never going to check it again and they wouldn't show up in the data.

MEMBER BREEN: I have a question for the developer in terms of your stratification based on ICU or floor. How are you attributing patients to each of those populations? Because as we just brought up, there are patients who obviously move around. Right? So, if a patient starts in the ICU, starts in the ED, goes to the ICU, is

discharged from the floor, how do you weight which part of their time? Which group are they going to fall into, in terms of their hyperglycemic scores, the ICU group or the floor group? How do you determine that?

MR. CAMPBELL: It is where they spend the majority of their time during that calendar day. So, if they spent the majority of the time of the ICU, then their day would contribute in the ICU score or that day would contribute for the ICU. And then once they transfer, the majority of the second day was in the acute care unit and it would contribute in that way.

MEMBER BREEN: Thank you.

MEMBER BAILEY: So my question comes along the lines of the end use. If the goal is for public reporting for hospitalcompare.gov or something similar, how do you deal with that stratification when you are reporting that to the public to consume?

Because we understand or if you

report it back to the institution and you stratify if by site of care or where the patients spent the preponderance of their time, the institution can understand that but how can the general consumer understand that when it is publicly reported?

MR. CAMPBELL: So, what we envision at this point is providing a score, for example, in easy to understand from the patient's perspective the measure score with critical care. So if I average percent of time for patients that were in critical care, you know, here is a score, and the average percentage of time in hyperglycemia for acute care. And then medical versus surgical patients.

So, that is, at this point, what we are considering. But a decision hasn't been made in terms of exactly how the measure would be reported yet on Hospital Compare or if it would be reported on Hospital Compare, I should say.

MEMBER BAILEY: So, a follow-up question on that would be a tertiary care hospital that gets very complex patients could appear to have very poor performance but a primary care center that doesn't have a high acuity may appear to do very well then. And wrong decisionmaking may occur from the patient perspective. Correct?

MR. CAMPBELL: Right. So, Almut, do you want to address that?

DR. WINTERSTEIN: Sure. Well, interestingly, since we have quite some diversity in our fuel testing hospitals, the tertiary care institution did actually very, very well. And that is probably really related to the fact that there was more of a knowledge base and how to deal with hyperglycemia on their standardized insulin infusion protocols available. And that was actually the same thing we saw when we were comparing our ICUs to the ACUs. The ICUs did actually better than the ACUs. And I think

that is the exact same reason we have standards protocols that start insulin drips in the OR for post-surgical patients and that insulin drip is continued in the ICU. Once patients are transferred to the ACU, things usually go a little bit more awry because then patients have to be switched to subcutaneous insulin. So, we actually saw the opposite than what would have been expected. And I think that is important to consider.

This is the nice part about having a surrogate outcome. It is actually quite manageable with insulin as long as someone has the knowhow to do so. And there are very good standardized insulin infusion protocols available these days to control glucose quite tightly in the ICU environment. The ACU is probably the bigger challenge.

CO-CHAIR GOLDEN: So I am going to just suggest that that was more of a usability question but we will get there.

Any other comments or questions on

validity? Bill?

MEMBER TAYLOR: I wonder if Tracy or one of our other experts who knows a lot about this could give us a little reflection on this. We are about to vote on specifications consistent with evidence. We heard from Jamie that the way the evidence was collected about the high glucose being associated with risk was collected in all sorts of ways that were different from this. And I would like to hear a little bit about how you put this together as we get to this vote.

DR. PACE: Can I also mention you are not just voting on the specifications consistent with the evidence. It is all that goes into validity. So, they also provided empirical validity testing, et cetera.

MEMBER BREEN: Yes, I think we are still going to have some more discussions when it comes to usability. But my take on this is we are talking about reliability and validity

testing of the process that they put forth to assess hyperglycemia, as they have defined greater than 200. It seems reasonable to me it has got good validity. It has got good reliability as a measure, to measure something. I think we already addressed the evidence of both those things. I think we still have some things to discuss about usability and how it might impact.

So, that is my take on how we -- it is a reasonable process.

CO-CHAIR GOLDEN: I think the issue is how it was collected. So waiting a day, not the ER, all those things go into those issues. But when somebody gets dinged, if you will, there is a reason and there is a reasonable reason for what I got dinged.

Ready to vote?

MS. BAL: Can people turn off their microphones? There we go. Okay, voting is open.

(Pause.)

MS. BAL: Okay, we have high four; moderate 14; low one. So, it will proceed.

CO-CHAIR GOLDEN: Feasibility.

CO-CHAIR ROSENZWEIG: I think here, I think that implementation of this measure would require a lot of effort and a lot of data collection but it is feasible, especially in large medical centers where they have the appropriate data management systems to be able to do this.

The issue was that in there are a lot of centers in which it would be very difficult to implement from the beginning but maybe the implementation, the initiation of a measure of this sort might lead to better data collection in the long-run.

CO-CHAIR GOLDEN: So, I have a question for the developer. In terms of case finding, I can understand if you are admitted with diabetes but now you are saying anybody with a blood sugar over 250. What kind of a mechanism or how difficult is it to run that

kind of data analysis, to identify those patients? And is that -- how burdensome is that?

MR. CAMPBELL: So this is Kyle Campbell from FMQAI. We did extensive feasibility testing with all of our field testing hospitals. And we found that the measure was feasible to extract and calculate without exception across all the hospitals. And we did submit as part of our submission NQF had recommended to provide a feasibility score card with regard to the measure. And I will note that the average score on that was a 2.85 across all 10:02:03.

So, in our findings, I think the folks on the Steering Committee are correct in saying that the hospital will have to have some effort in terms of setting the measure up but it will be much less burdensome for the hospital, once the measure is programmed compared to manual chart of distraction.

And so for that reason and with

the reason that we found that all of these data were retrievable, including the point of care data that was discussed earlier, we consider this measure from an electronic perspective highly feasible.

CO-CHAIR GOLDEN: And a follow-up question. When you were doing your testing, how much software programming needed to be done? And were the test hospitals reimbursed for the software programming?

MR. CAMPBELL: Actually in this case, the test hospitals were not reimbursed for the software programming. They volunteered to participate in this project with us.

I would say on average -- I don't know. Maybe I will defer this question Almut because she was closely involved at the academic institution where we did our formative testing. Do you have a sense, Almut, of how many hours in terms of programming?

DR. WINTERSTEIN: I think it is really important to differentiate between the initial setup and any subsequent data retrieval.

And I hope that I don't go too long now but obviously one of the most important data elements in here is lab values. And just to give you an idea about this, our hospital has about 800 different lab values in the system and different ways those can be ordered. And if you look at glucose, glucose can be ordered as part of several test batteries.

So to find the individual orders that could have been a metabolic panel or just a normal daily morning draw and so forth, as well as the chem sticks of course, there is actually a variety of different lab values that can be used.

Now, as part of meaningful use, lab values will have to be standardized and coded with coding systems, which eventually

will look like an ICD-9 code. So, in a year from now very likely most hospitals can just be told we want code 1, 2, 3, 4, 5 and that will automatically pull all of those glucose values.

In our case, this hadn't been implemented as it hadn't for many hospitals, which basically means there needs to be a hand search of the lab values that really need to be included into the measure. Once this is done, the extraction out of the system is very simple and straight forward. So, it is more really using the dictionary correctly to find the information that is really needed that concerns the medications as well as the laboratory values that were used in this particular measure. All other pieces before you are standardized already because they actually belong to the charge or the administrative systems that are use and hospitals and that have been used for many years. So, admission tests and things like

that are in fact standardized data values already.

So altogether, I think once the measure is set up, the retrieval itself is essentially really a push of a button. And the code that we wrote to run the measure of values that looks at the admission time and defines when patients enter the denominator and numerator and so forth, all of this readily available.

CO-CHAIR GOLDEN: Comments or questions?

DR. BURSTIN: I just wanted to point out this is actually one of the first de novo eMeasures that NQF has received. We did this work for CMS and ONC about a year ago, I guess, to come up with a score about data feasibility because there was a lot of concerns about new eMeasures being brought forward where the data was actually not feasible to collect. So, these were actually done with an expert panel, came up with those

particular data availability, data accuracy and data standards, to allow committees like you to have some confidence that what you are putting out there is actually findable and in a standardized and reliable way in an EHR.

So, this is actually quite a good score overall but actually we are just really pleased to sort of see the light of day. And I think we have heard a lot, particularly from hospitals about how difficult it is, sometimes, to implement some of the eMeasures.

So, I think since this is a relatively not very complex set of data to put together, but just wanted to put that -- give you that perspective.

CO-CHAIR GOLDEN: Any other comments and questions? Ready to vote?

MS. BAL: Voting is open.

(Pause.)

MS. BAL: Okay, we have high nine; moderate eight; low one.

CO-CHAIR GOLDEN: Usability.

CO-CHAIR ROSENZWEIG: We didn't discuss this at length. I don't think we had time to discuss this at length at our conference call. But my sense is that this is a usable measure.

CO-CHAIR GOLDEN: I guess a question for the developers in terms of CMS, not every hospital can do this. What is your perception over time about requiring reporting versus -- I mean obviously if you have an EMR you can do it but if you don't have an EMR, you can't.

So, are there plans down the road about requiring hospitals to do this over a period of time?

MR. CAMPBELL: This is the developer. And I would not be able to answer that question for CMS. I don't know if anyone from CMS is available to respond.

MS. BODKIN: Hi, I'm Noni Bodkin. I am from CMS. I am from the Quality Measure Health Assessment Group in the Centers for

Clinical Standards and Quality. So, that is the birthplace of many, if not most of the quality measures you are familiar with.

And I cannot speak about our policy but it is my understanding that the first step for this measure is to be considered for our meaningful use Stage III.

So, that is a really important initial step. We don't know what will happen. We are very excited to have probably the first de novo ECQM before a Steering Committee.

So, we will keep you posted on that and we will follow all the normal processes with public comment, our notice of public rulemaking, and very careful evaluation. Thank you.

MEMBER McCOLLISTER-SLIPP: So, I have a question about that because I mean is this something -- so, this is going to require some degree of algorithm using EHR data of some form or another. And one of the questions I had and I didn't know if it was

around harmonization that we would get to this is how does this fit into meaningful use. And I mean I understand with where we are with meaningful use but I haven't gone through and looked at the specific elements of it.

So, is this data not the composite measure, which is calculated by data that is entered. Is all this stuff part of meaningful use as it currently stands? Do you know?

CO-CHAIR GOLDEN: I think meaningful use here would be that this measure would be a requirement to achieve meaningful use Stage III certification. You have to be able to calculate this measure.

MEMBER McCOLLISTER-SLIPP: Well, that is Stage III but is the data, are the data fields that as currently dictated part of the existing like meaningful use Stage I and II in terms of having those fields already available?

CO-CHAIR GOLDEN: I can tell you from experience of looking at my own

institutional EMR, lots of data fields exist. Whether you can collect them all is another matter altogether. So, just because the data fields exist doesn't mean they can be integrated. It gets, unfortunately, it is often a disappointing level of why isn't this working better.

Bob?

MEMBER BAILEY: So, one could argue, since we are at use and usability now, that incorporating it into meaningful use, there are financial incentives for institutions to implement these programs. And, as Tracy mentioned and as Sue mentioned earlier, there are opportunities for these organizations to look at their data, determine where they are, then also determine where they are with respect to their peers.

So, I see this as benefits. And then if institutions don't comply with meaningful use Stage III, subsequently they will be financially penalized. So, it is a

positive outcome, I would suggest.

CO-CHAIR GOLDEN: Tracy.

MEMBER BREEN: I think it is interesting because there is nothing there right now. I mean this is truly a new way for hospitals to look at data that they have been sitting on for many, many years.

I think that it is very usable. It will be interesting to see and it is a place to start, in my mind. My reservations are so small in terms of the unintended negative consequences, to give one anecdote, though about unintended consequences, one of the SCIP scores used in the post-operative CTICU realm has been no patient to have a glucose greater than 200 without a lower level. So, zero is less than 200, as I like to say. And so one of the unintended consequences has been that many CTICUs or post-operative units have done very well on their SCIP scores, they don't have any glucoses higher than 200. But then when you

begin to look at their hypoglycemia, we are going to get to hypoglycemia, they have had surprisingly high hypoglycemia that they didn't know because they weren't looking at it.

So, I think the power of looking at something is very powerful. So, I think this tool will be a very powerful tool for hospitals and I think some people are going to get a very rude awakening of what their numbers are.

I think it is reassuring to me that they are going to be at least shown to other similar hospitals. I think the stratification of place seems reasonable, ICU versus the floor. And again, it is a good first start. My only tiny reservation is that again, how we are defining persistent hyperglycemia by basically putting this in play, does it then prevent other potentially more robust ways of measuring hyperglycemia from being investigated? Whatever. Smart

people out there will figure out what to do. But that is a very, very small revision.

CO-CHAIR ROSENZWEIG: You know I also think, though, that the process of implementing this would allow for collection of a lot of data, much more data than we currently have. And that whatever the specific thresholds for defining good versus bad control in the hospital are identified here, it could lead to more accurate data in the future.

CO-CHAIR GOLDEN: Are we ready to vote?

MS. BAL: Voting is open.

(Pause.)

MS. BAL: So, we have high 11; moderate seven; low one.

CO-CHAIR GOLDEN: All right, now we are into the big picture. So any final comments on this, otherwise we can vote on endorsement, yes or no.

Not yet. Not yet. We have a

comment from Sue.

MEMBER KIRKMAN: Just to clarify. This is going to be paired with the hypoglycemia, right? So what she mentioned about people could look really good on this measure because their patients are all in the 30s, potentially wouldn't happen. Right?

CO-CHAIR GOLDEN: Yes, you know pairing is interesting in that they are not necessarily -- well, yes, they won't be in isolation.

MEMBER BAILEY: So just to be clear, so if this one passes and the hypoglycemia one does not, then what happens?

CO-CHAIR GOLDEN: We will think about that.

(Laughter.)

MEMBER McCOLLISTER-SLIPP: Just for color commentary, my father has Type 2 diabetes. He was diagnosed, I don't know, maybe about 20 years ago. Al the complications. He is in his 80s. And he has

had lots of hospital visits over the last few years. Because I am nerdy enough to know about the data about in-hospital glucose control, I have had my mother stay on top of it.

In terms of like the hospitals variability in different institutions in the region where he is, it is considerable in terms of their willingness, and their ability, and their intent to focus on this particular measure. I mean some of them are completely on top of it, some they would never check if my mother wasn't bothering them or if he didn't have his own meter.

So, I think this is a very helpful important step for us to do to incentivize.

CO-CHAIR GOLDEN: Let's open the polls.

MS. BAL: Voting is open.

(Pause.)

MS. BAL: Okay, we have yes, 18 and no, one.

CO-CHAIR GOLDEN: All right. Thank you all. I think that the next one will probably go a little faster.

So, it is 10:20. So I would suggest we reset our biologic parameters and re-gather at 10:30. So take a break.

(Whereupon, the foregoing matter went off the record at 10:16 a.m. and went back on the record at 10:29 a.m.)

CO-CHAIR GOLDEN: So, as we being to re-gather, before we do the next measure, it is time for I think if we are going to be doing a variation on the Powerball. And yesterday people were talking about the fact that you would be open for a two- or a three-year term. And so there is some -- do we have the -- where is the magic hat? Oh, there it is. Okay.

So we have the sippy cup and people will choose their fate. It is like a Magic 8 ball here.

So we are going to be going around and let people know what your number is.

MS. TIGHE: And I will just jump in with a little bit of the rational for this. We are moving to standing committees, as we discussed yesterday. And so we do need to kind of split your term, so we are not refilling the committee in its entirety all at one point in time. That said, if you get a two-year term and you are just really wanting to stay on the committee for another three-year term, you are welcome to reapply at that point in time. Our policy is written so that committee members can stay on for two terms in a row. After that, we will ask you to stay off for a full term before applying again.

So, two years could be extended to five if you wanted it to. Three years could be extended to six, if you wanted it to. Either way we have appreciated having you on at least for this meeting and two years going forward.

CO-CHAIR ROSENZWEIG: How about allowing people to trade terms?

(Laughter.)

MS. TIGHE: Certainly, if there are any conflict with what you draw or need to reevaluate the terms, just let our staff know.

CO-CHAIR GOLDEN: So, you have these numbers and do you have to have folks report to you about the numbers or how do you want to handle these?

MS. TIGHE: Poonam will go around and record.

CO-CHAIR GOLDEN: Okay, good.

CO-CHAIR ROSENZWEIG: I'll go ahead to the next measure, which is glycemic control hypoglycemia. Let me just get my materials here.

This is number 2363. And are the measure developers on the line? So, if you would like to describe the measures to us, we would appreciate it.

DR. WINTERSTEIN: Yes, we are

here. I have to unmute myself again.

MR. CAMPBELL: Yes, I am here as well.

MS. TIGHE: Kyle, do you have any comments to add about this measure before the committee begins to review?

MR. CAMPBELL: You know in the beginning when we introduced this, we introduced both of them at the same time. I would just say that for public reporting, we are looking at severe hypoglycemic events, so then less than or equal to 40 milligrams per deciliter. And this threshold has been consistent with a definition of safety in the majority of clinical trials. And we found, through our work in testing these measures that most of these events, at least according to our clinical reviews, largely prevent it.

So, thank you and we appreciate your consideration and look forward to any questions you have about this measure.

CO-CHAIR GOLDEN: Okay.

MEMBER LEDDY: We can start.

CO-CHAIR GOLDEN: Thank you.

MEMBER LEDDY: This measure looks at the rate of hypoglycemic events following the administration of an anti-diabetic agent. The rationale relates to glycemic control and hypoglycemia management in the hospital inpatient setting and is proposed as a companion measure to the measure we just discussed and voted on, glycemic control.

It is an intermediate outcome occurring in the inpatient setting. It has serious consequences, longer stays in the hospital, increased mortality.

So, the developers see important benefits arising from the implementation of this measure. One is that providers will be incentivized to recognize hypoglycemia, prevent it, treat it, and as a result, there will be shorter hospital stays and fewer or lower mortality.

It does really point toward

advancing quality of care in this area of patient safety. It has been identified as a very important issue by the National Quality Strategy.

The numerator statement looks at the total number of hypoglycemic events that were preceded by the administration of rapid or short-acting insulin within 12 hours or an anti-diabetic agent, other than short-acting insulin administered within 24 hours, not followed by another measure of glucose greater than 50 within five minutes. And the events would need to be at least 20 hours apart.

There was an optional numerator statement given but the developer just mentioned that the key number here is less than 40. That is a very critical glucose value.

And let's see. The denominator's total number of hospital days on which an anti-diabetic agent was administered. There is a single denominator exclusion, which is

admissions less than -- admissions greater than 120 days. It is my understanding that this 120 day definition is really standard for looking at a lot of measures.

This is an outcome measure. The data is from electronic records. Electronic clinical data in the pharmacy and the laboratory. It is a facility-level analysis.

So moving on to evidence, I think the evidence is very good. Five studies were reported. Guideline recommendations were included. However, these were not -- the guidelines that were included were not the focus of the measure. They had to do with treating hyperglycemia. Of course, this measure is about treating hypoglycemia or preventing it. The evidence, however, is sufficient.

The main points are that patients with hypoglycemia, less than 40, are at greater risk of in-hospital mortality and that patients with hypoglycemia had longer lengths

of stay.

We did not use a grading system but the studies were described carefully and the results were consistent.

There are two ways of looking at the evidence. One, if one believes that it represents a formal systematic review, one goes to the algorithm boxes 3, 4, and 5b. If not, you move down to 7, 8, and 9. But either way, we get to a moderate rating.

So perhaps this would be a good time to talk about the evidence. Any comments?

We all know that hypoglycemia is very dangerous and the hospital is a very dangerous place for Type 1 diabetics and to a little bit lesser extent, Type 2s. You have to have a family member there to take care of them.

CO-CHAIR ROSENZWEIG: Is it clear really that hypoglycemia should be considered in an intermediate outcome? I mean in and of

itself, isn't it a --

MEMBER LEDDY: That is a really good point because it could be curtains --

CO-CHAIR ROSENZWEIG: Exactly.

MEMBER LEDDY: -- could be the outcome, the ultimate outcome.

CO-CHAIR ROSENZWEIG: So I mean it is intermediate towards death or brain damage. But seriously, I think it could be considered an actual negative outcome in and of itself.

MEMBER McCOLLISTER-SLIPP: Sorry. I need to hold my card up or whatever. But would that make a difference in terms of the implications of this measure? I mean so whether ti is intermediate or just an outcome?

DR. BURSTIN: The only difference and Karen can weigh in on this is that if it is an outcome, you just have to provide a rationale for the evidence, as opposed to firm evidence because there is just sort of on the very basis, a bad outcome is a bad outcome. How much evidence do you need for it?

Anything to add?

DR. PACE: Though this would probably be considered an intermediate.

DR. BURSTIN: But they are just saying now they think it is an outcome, not an intermediate.

MEMBER LEDDY: Intermediate outcome.

DR. PACE: Right. And I think this is a problem with our submission form. I don't know that we have that option for an intermediate outcome.

So for intermediate outcome, we like to see the link to other outcomes. We have talked about that, mortality and morbidity. If it is an end result outcome, then as Helen said, we ask for a rationale that there are healthcare services and interventions that can affect it but it is really more of a rationale discussion than the actual evidence.

CO-CHAIR ROSENZWEIG: Because the

staff review had mentioned that it is an intermediate outcome and requires evidence of linkage to health outcomes.

DR. PACE: Right. So, you think it is more of an outcome. Okay, that's fine.

CO-CHAIR ROSENZWEIG: Yes, Sue?

MEMBER KIRKMAN: So, I mean it is potentially an outcome of treatment but it also is a marker for a sicker patient. So, I think in some ways it is a little bit of an intermediate outcome. It is a little bit semantics. But even people not on insulin can have severe hypoglycemia in the hospital and they have a very poor prognosis.

So, there is linkage to the final outcome but it is not necessarily --

DR. PACE: So you know this where things stop and outcome begins, there is not hard and fast lines. I think maybe we can hear from the developer. I think they submitted evidence of linkage to other outcomes. So, they probably conceptualized it

that way. Developer?

MR. CAMPBELL: Yes, this is Kyle Campbell at FMQAI. We did envision it as an intermediate outcome with a hard outcome of mortality and morbidity. And as mentioned, we identified, I think, nine in all now with the supplement that we provided to the Steering Committee, you know evaluated mortality. And there was consistency in magnitude of direction and effect for those studies.

I mean my take on this is we would still consider it an intermediate outcome, in keeping with the nomenclature for the hyperglycemia measure.

DR. PACE: Is there any question of the linkage between this intermediate outcome and mortality and other morbidity?

CO-CHAIR ROSENZWEIG: I think there is. Well, our review work addressed that, yes.

MEMBER LEDDY: I think there is really strong linkage between the evidence and

the outcomes.

CO-CHAIR ROSENZWEIG: I think one issue that should be mentioned is that the evidence between linkage of hypoglycemic events and falls is reasonably robust. And falls are a never event for payment for hospitals. So, they are very, very incentivizes to reduce hypoglycemia. I think there was some discussion as to whether or not severe hypoglycemia itself was to be considered a never event. I don't know if that has happened or not. Does anyone?

But clearly, the incidence of a fall, of major falls have to be reported by hospitals and the causes of them have to be determined. And in patients with hypoglycemia, this could be --

MS. TIGHE: I worked on the NQF Serious Reportable Events Report and I will just speak to the fact that it is classified under a medication error because typically you have that severe hypoglycemia as a result of

mismanagement of medication. So, it is something captured as an NQF serious reportable event, which is the foundation for the never events.

CO-CHAIR ROSENZWEIG: So you have to actually determine that there is a mismanagement of medication?

MS. TIGHE: The trigger for the event is patient injury or serious harm. But yes, it is related to the medication mismanagement.

CO-CHAIR ROSENZWEIG: Well, in any case, clearly the hospitals that implement this would be very highly incentivized to measure this.

So any other comments?

MEMBER LEDDY: I have no other comments.

CO-CHAIR ROSENZWEIG: Any questions?

MEMBER LEDDY: Can we vote?

CO-CHAIR ROSENZWEIG: I think we

can, yes.

MS. BAL: Voting is open.

(Pause.)

MS. BAL: Okay, high 13; moderate six.

MEMBER LEDDY: Okay, moving on to 1b, gap in care. This was a new measure. So, it was presented to eight testing hospitals. The performance data is available. The range of occurrence of or reported performance was between 0.36 percent and 0.89 percent.

This is really when you look at things in a low incidence of outcome but considering that it is such a severe and dangerous event, even those low numbers look like a gap to me.

Looking at the cited studies, there was no address made to the variability and performance but the rates for patient days were higher in patients in the ICU than in the non-ICU setting. And then some rates per admissions were between 2.3 and 3.5 percent.

So I think that even though it is quite small, really any gap is very serious. And I would like to call for a vote, except Bob has a comment.

MEMBER BAILEY: Although there is not much of a gap now, we need to consider that there may be a widening gap going forward with a focus on controlling hyperglycemia. So, that should come into consideration as well.

MEMBER LEDDY: Absolutely. It is so important to have this as a companion measure. A very, very good point.

CO-CHAIR ROSENZWEIG: Okay, let's vote.

MS. BAL: Okay, voting is open.

(Pause.)

MS. BAL: Okay, we have high 12, moderate six; low one.

CO-CHAIR ROSENZWEIG: Okay, going on to --

MEMBER LEDDY: To 1b, disparities.

The developer actually stratified the measure by age, race, ethnicity and payer source and concluded there were no significant differences by race or age groups.

There was no information on how it was determined that there were no significant differences.

The rates varied by race for the Hispanics, 0.4 percent; African Americans, 0.67 percent.

There were a couple of studies cited. I think one of them was in an acute MI setting, where when they looked at incidence of hypoglycemia, it was more common in elderly people, women and one ethnic group, which I don't remember. And then there was another study cited where there was also older patients and also women.

But to my review, it just did not seem to me that disparities was really an issue.

MS. TIGHE: Just a process point.

That actually falls under 1b, which we just voted on. Not to cut conversation short.

MEMBER LEDDY: Okay, thank you for your help.

MS. TIGHE: That shouldn't be factored into your vote on priority of high impact.

MEMBER LEDDY: Okay, so here we are at priority. The occurrence of hypoglycemia in the hospital is a very significant adverse drug reaction. It is the most common drug reaction and accounts for about a third of all adverse drug reactions in the hospital. It is a very high priority for I guess the National Quality Survey.

So, does the measure address a significant health problem? And may we vote?

MS. BAL: Voting is open.

(Pause.)

MS. BAL: Okay, the results are high 17; moderate two.

CO-CHAIR ROSENZWEIG: Okay,

reliability. Any comments?

MEMBER LEDDY: We have discussed in the previous measure the eMeasure. And the eMeasure technical review that has been. I believe the developers have given us a good reason not to consider the optional numerator. It would, if we had a higher optional numerator of less than 70, maybe more folks would have their daily glucose records reviewed in the hospital by their provider and maybe some more hypoglycemic events of severity would be avoided. But I, personally, would like to see it remain at 40 or less.

I would remove the optional numerator. Any comments about that?

MR. CAMPBELL: This is the measure developer. And just to comment in terms of the optional numerator, when this measure went out for national public comment, we received comments from hospitals that asked us to consider for internal quality improvement an optional numerator. And so that is the origin

of why the optional numerator was included.

We are not recommending that for public reporting but internal quality improvement only.

MEMBER LEDDY: Is it possible to have two numerators in the same measure?

MR. CAMPBELL: It is from our perspective because the publicly reported measure would just be one numerator.

DR. PACE: I think that would need to be clearly specified. The way it is now, it makes it look like you have the option of using whichever one you want. So, we need to be clear what NQF endorsement would be.

CO-CHAIR GOLDEN: So it gets back to our discussion yesterday with NCQA.

Would CMS be adverse to eliminating 70?

MR. CAMPBELL: We would not.

CO-CHAIR ROSENZWEIG: One question I would have related also is usually hypoglycemia is often defined, severe

hypoglycemia is often defined as requiring help by a healthcare provider or by another person in order to reverse it. Or there may be issues related to whether or not there is loss of consciousness involved with this or Whipple's Triad. And none of those are mentioned. This is purely just looking at the number of actual times in which blood glucose less than 40 is listed.

Did the measure developers consider issues related to that?

MR. CAMPBELL: This is Kyle Campbell again from FMQAI. We did consider that but our focus was on what was readily extractable from electronic health record. And that type of information would necessitate chart review in order to include.

CO-CHAIR GOLDEN: Yes, Sue?

MEMBER KIRKMAN: So, two things. Just a comment on yours. I mean I think for the inpatient setting, it is difficult because if someone is comatose and their blood glucose

is 69, a health care professional would treat it anyway. So it becomes a little difficult.

But my question goes back to the optional numerator. So I mean I do think for internal quality improvement, you wouldn't want to just look at less than 40 because that is going to be vanishingly rare in your hospital, hopefully. But if you have a lot of people in the 50s and 60s, then that is a problem that you would probably want to address internally. So, I guess it is more of a process question for NQF. I mean, is there a possibility to endorse the publicly reported measure with an optional for internal QA or is that just not -- it has to be all or none and anything we endorse could be publicly reported. Is that the situation?

DR. PACE: I think you have some options here but I just think it has to be very clear in the specifications that component, it is not really optional. It is another way to -- it is not optional in terms

of how you implement the measure for an accountability purpose, if that is what you all are saying.

I mean you have some options here in terms of endorsing it with kind of reporting both, endorse the less than 40 as the one that is used in accountability applications and it being very clear that the less than 70 is the QI component or ask them to remove it.

So, I think you have to have some discussion about that.

DR. BURSTIN: It is also worth noting, I believe these measures and notings here have already been proposed as part of some of the ongoing upcoming programs for CMS. So, it may also be that you may want, depending on one denominator there may just be notes to implementers that you could something along the lines of the optional.

But I don't know if anyone has any comment about that. Too early. So again, it

could be something we could ask them to consider and come back to us, just to keep it simple. But for now, I would assume you should look at what is the actual specification, not the optional one. And we can allow them to come back to us with some thoughts about how we handle the optional piece of it.

But since this measure may likely be on some public reporting or payment program in the future, I think we need to focus on the accountability application, at this point.

MEMBER KIRKMAN: So does everything that is endorsed by NQF could potentially be publicly reported? Is that -- okay.

CO-CHAIR GOLDEN: Yes, I am a little concerned by that if we make it optional, it means that CMS has the option of making it 70, which I don't think --

DR. BURSTIN: No. So the endorsed measure would be whatever you think it should

be. I think then the question is how does CMS want to handle potentially asking hospitals to submit the other data more for learning than anything else, as opposed to that.

But again, I think we need to get some other input from CMS.

DR. PACE: Right. And perhaps one way to handle it is Helen was talking about, maybe rather than in the specifications, it is handled under potential uses, where that may be more comfortable than having it right in the specifications.

So, I think you should first consider the accountability application and then make recommendations about how to handle the --

MEMBER KIRKMAN: Yes, I just wouldn't want the message that hospitals get that anything above 40 is okay. I mean I just don't know how to handle that.

MEMBER BREEN: It sounds like if they have language in their comment to the

hospitals about how to use it, then it could be there.

I agree with Anne, though, for the purposes of this, it seems like it would be clearer to just vote on this without the optional piece, even though we all know that the hospitals should be doing that. But again, it seems like it would blur lines to make it murky if it looks like that was reportable.

CO-CHAIR ROSENZWEIG: With respect to the specifications, does this measure allow for distinguishing between whether or not the anti-hyperglycemic medication was administered intentionally or according to the orders or by mistake? Since some of these episodes occur because -- I think probably a significant number occur because of problems with the mistakes made.

MR. CAMPBELL: So this is the measure developer. We don't distinguish the administration of the medications by mistake.

But information comes from the electronic, the eMAR, the electronic medication administration record. So, we do have data that the drug was actually administered.

CO-CHAIR ROSENZWEIG: Okay, I'm not suggesting that it has to be but it would be of interest.

MEMBER McCOLLISTER-SLIPP: So, a couple points to your point earlier about unconscious hypos or adverse events, I mean that is highly variable from patient to patient. So, I think relying on adverse events or being unconscious or something really bad happening is really, it is just better to have a number.

I would think that 40 is a little low. That makes me a little nervous. I think 70 may be a little high. So, I don't know what the sweet spot is but the 40 is really low.

And then when it comes to again, meter variability, and again, I am more

familiar with consumer outpatient meters than I am with point of care meters but once you get that low, you get into reliability issues. So, I think we need to think about that. I mean, is 40 really 40 or is it 20 or is it 50?

And that is our job is not to determine whether or not these meters are accurate but I mean a 40 would make me very nervous and I think we need to have an incentive for hospitals to not --

I mean this is something I am pretty sensitive to because I am pretty insulin-sensitive when I have been hospitalized, which fortunately hasn't been for a while. They always want to overdose me and I refuse the medication because inevitably they will look at my blood sugar, which is high and they will try to give me 10 units of insulin and I am just like no, I don't want to die.

So, I mean there needs to be an incentive for hospitals to be very careful

about that, not that I think that people are being careless or being egregious in their carelessness. But anyway, I think I am sort of rambling at this point. But this is a significant issue and I think we need to give real thought to does it make sense to have that flexible numerator and is that area, is that wide variability of that number an appropriate range.

CO-CHAIR ROSENZWEIG: Tracy.

MEMBER BREEN: You know where to draw your line in the sand on the hypo front is very similar to the conversation we just had on where to draw the line on the hyper front. At some point, it is just high. All right, so I think just like they had consensus around anything greater than 200, most people are going to agree that that is really high, drawing the line at 40, everyone agrees that that is low.

As you begin to creeping up higher, this comes out to discussions with

laboratory people. How do you define a critical value and where you draw that line.

But 40 is interesting because when you look at meter variability and what is an accepted range of what is accurate, when you match it up to labs, unfortunately, you are allowed kind of a 10 to 15 point variability in your meters.

So, if you start doing the math on that, if most of these glucose checks are done by point of care meters and you come up with a glucose check of 50 on the meter, it might actually be 65 or 70. It could also be low. And then you get into a glucose of 60 is very normal for many people. In pregnancy, we actually want that. Right? So, there is a whole host of glucoses of 55 to 60 that are very normal and healthy for many people in the hospital.

At 40 though, even when you do that math, maybe their glucose is really 55, that is still really low. So for me, a

glucose of 40 is a nice anchor for just a really severe hypoglycemic event. I think there will be very few false negatives in that pool and at least it gives the hospital some kind of anchor as to what is really bad.

The other question is well, how do you get not just the really bad but the not so great. I think that is the other discussion we had about the optional numerator. But again, considering there is nothing right there now, to me it seems like a place to start, just like the hyperglycemic reporting was a place to start.

MEMBER McCOLLISTER-SLIPP: But the consequences of being slightly off for hyperglycemia is very different than the consequences for being slightly off for hypoglycemia. I mean, especially if it somebody is Type 2 and has lots of risk factors for -- or Type 1 and has lots or risk factors for cardiovascular disease.

MEMBER BREEN: Yes, but I think we

are saying that this is a reportable measure. You know, it gives a hospital a way to anchor themselves vis-a-vis other hospitals of how they are doing with this severe hyperglycemia. So someone doing very badly with severe hyperglycemia is also probably not doing so great with a not so severe hypoglycemia. To me, it is just a marker for where you are, vis-a-vis other hospitals. There is no -- I don't think we are ever going to find a number that everyone agrees on. When you look at the literature on severe hypoglycemia, they even define it differently, how you define severe hypoglycemia. It is kind of a made up term.

But I think most groups would agree that 40 is bad. So once you get above that --

MEMBER McCOLLISTER-SLIPP: For verification, is this hypo or is this severe hypo?

MEMBER BREEN: They are calling it -- so now we get into -- they are calling it

hypoglycemia but this is really severe hypoglycemia. Hypoglycemia we know is defined as less than 70. Right? If you look at all the ADA criteria, hypoglycemia is defined as less than 70 but that becomes problematic in non-diabetic patient population or in a pregnant patient population with diabetes.

So, there may be some nomenclature issues about what we suggest to the developer in terms of are we calling this hypoglycemia or severe hypoglycemia. That is a really good point.

MEMBER McCOLLISTER-SLIPP: Don't we already have existing issues around severe hypo though? I mean that is one of the few things that is actually reportable to FDA through AERS. I mean I know that this is inpatient and it is completely different. But I would think that we would -- there is a distinction between severe hypo, which hopefully is a never event, and hypo.

CO-CHAIR GOLDEN: A little while

ago, I looked up never events and it is severe hypoglycemia with injury or morbidity. So, it actually requires some adverse consequence, which this would not.

MEMBER BREEN: This is potentially a bigger pool than that. Right? It seems like we are expanding the pool and the concept of medication error is a huge one but not for this setting, how we report errors.

And I would also just comment in terms of how hospitals currently look at hypoglycemia, this measure to me is very clinically interesting because it makes it patient-specific or at least it relates it to patients on high-risk medications. Many hospitals right now, when they try to look at their hypoglycemia data just go on the ACCU-CHEK data or they go so they take their gazillion point of care measurements and they say what percent of those gazillion point of care measurements are hypoglycemia. And they feel really good about themselves because it

is like 0.5 percent or one percent. It is not clinically valid. But this seems to be a much more clinically relevant way for hospital to look out of all the patients treated, out of all the episodes that someone received a treatment with a high-risk medication, how many of those were associated with severe hypoglycemia, whether or not it related to harm? That gives you a bigger piece of the pie to kind of think about.

CO-CHAIR GOLDEN: Sue.

MEMBER KIRKMAN: So again, I think this gets back to the issue of just making sure that hospitals don't think that less than 40 is all they have to worry about. Because I agree with you. I don't think less than 70 should be publicly reported, necessarily. But I just really hope that however the specifications are set up and however people are doing their internal QA, they don't realize that there is a lot of other hypoglycemia. I mean, if you are not going to

call this severe hypoglycemia -- I mean that would be one thing would be to call this severe hypoglycemia and say that is what is publicly reported. Because again, by saying this is a hypoglycemia measurement and hypoglycemia is less than 40, then that sort of implies that anything between 40 and 70 is high.

MEMBER BREEN: Because it is not actually.

MEMBER KIRKMAN: I mean it is fine.

MEMBER BREEN: It is not the right definition. If we are saying hypoglycemia, this is not the definition of hypoglycemia.

MEMBER KIRKMAN: Yes, either call it severe hypoglycemia or the specifications just have to be really clear or the implementation has to be really clear that hospitals can't ignore anything from 41 to 70.

DR. PACE: So I guess maybe we can hear from the developer again if they would be

amenable to labeling it severe hypoglycemia. And I think you have already said that you would be willing to take the optional out of the specification. So maybe you can talk about it under improvement.

MEMBER McCOLLISTER-SLIPP: But I think it is appropriate to incentivize hospitals to not keep their patients hypo. I mean we don't want to go hyper but that is an incredibly uncomfortable place to hang out.

And again, I don't think there are people out there who would sort of intentionally keep their patients at a lower level but I think it is appropriate to distinguish hypo from severe hypo.

So I mean the question is, is what is that measure? And there is significant variability from patient to patient in terms of the consequences of that measure. I mean, if you are coming from a high and you are at 70, you may have heart palpitations and sweat. I mean there could be -- you could even go

unconscious. I know people who don't have great control who become unconscious at 70.

So, I mean in think it is important to have a distinction between the two, generally speaking. So, I think it is appropriate to have a hypo measure but 40 makes me nervous just because that is really low, no matter how you define it.

DR. PACE: So, you are suggesting to have a measure with basically both measures; less than 70 and less than 40? Is that what you are advocating?

MEMBER McCOLLISTER-SLIPP: Well, I think we need to hear from the developer. It would be nice to hear from the developer about what their goal of this is. If it is really severe hypoglycemia, then maybe we just need to name it.

MR. CAMPBELL: This is Kyle Campbell again from FMQAI and that is our goal, is severe hypoglycemia. From an accountability perspective, we did review this

with our expert panel and we found that those events that are less than 40 are definitely preventable. And our concern with a publicly reported measure, where the measure would be let's say less than 70, that we begin to see that many of those values aren't preventable from the hospital side.

So while we do acknowledge that we think hospitals from an internal QI perspective should be looking at less than 70 as a mild or hypoglycemia metric for internal QI, for the reporting function, we would be amenable to relabeling this as severe hypoglycemia.

CO-CHAIR GOLDEN: Just to try to shape this, the thresholding is almost a validity question, not a reliability question. The question is can we -- is it defined so we have a number. And can you measure that in a reliable way?

DR. PACE: And they did present some reliability testing data.

CO-CHAIR GOLDEN: Yes. So, whether it is 50 or 60 or 40, the question is, can you collect that data? And then we just repeat the same thing for validity. But I think we are discussing validity here and not reliability.

MS. BAL: Are we voting?

CO-CHAIR GOLDEN: I think so. Yes, we are voting on this.

MS. BAL: Okay, go ahead and vote.

(Pause.)

MS. BAL: Okay, we have high 11; moderate seven; low one.

CO-CHAIR ROSENZWEIG: Okay, comments about validity?

Yes, Patricia.

MEMBER McDERMOTT: I've been trying to figure out when to fit this in.

In the denominator exclusion says admissions with length of stay greater than 120 days. This is supposed to be a paired measure with hyperglycemia, which had anything

greater than ten days or after days excluded. How does this correlate?

MR. CAMPBELL: This is the measure developer. The exclusion for greater than 120 days, I believe, one of the Steering Committee members brought this up earlier is that has been an exclusion of part of the inpatient quality reporting measures by CMS and we adopted it.

There are very few, and in fact in this measure I need to look, but there are very few, if any, patients that are excluded due to that exclusion.

With regard to the hyperglycemia, we are only looking at the first ten days of measurement as Almut suggested because we are trying to avoid long-term stay patients skewing the measure results. So, if the patient was actually in the hospital for 90 days and contributed a large number of hyperglycemic days, we didn't want to skew the measure results with that measurement. So

that is the rationale here.

CO-CHAIR GOLDEN: But I think also keep in mind the first measure was a percent. So, if you had a long stay, you can dilute the percentage. This is a total number, correct number of events.

Though, to follow up though, if you have one very brittle patient, will that skew your numbers as well? Have you given any thought to the fact that if you have one patient that jumps around, could that one person be -- I mean and whether that is relevant or not to be discussed by the group.

MR. CAMPBELL: I am going to defer that question to Almut, who worked with our technical expert panel workgroup on that issue.

DR. WINTERSTEIN: Yes, just to follow-up what Kyle just said. So patients in the hypoglycemia measure, patients is the length of stay more than ten days are not excluded from the measure. We just truncate

the follow-up at this point. So, the patient populations that are included in terms of their length of stay are actually identical in both measures.

In terms of having patients contribute more or less in the hypoglycemia measures, based on the number of events they have, this is an incident-based measure. So we are counting the total number of events over the total number of days that patients contributed in the hospital.

And the reason we did this is because we noticed that there actually are some patients who have repeated incidents of hypoglycemia. And we didn't want to take this out because it seemed to be important enough to capture every single event.

So, what we are essentially counting right now is the total number of events that occurred during admission. Those events have to be at least 20 hours apart to assure that we are not looking at the same

incidence because there were remeasurements done. And then we normalized this, we standardized this to the total amount of time that was available.

CO-CHAIR ROSENZWEIG: Sue?

MEMBER KIRKMAN: Yes, to me it seems reasonable because this is almost a safety issue. If someone is in the hospital 80 days and they get hypoglycemic 60 times, I mean I think that is not really skewing the data. That is really bad care.

Whereas, the hyperglycemia measure, the way it is calculated, I can see that your hospital's quality of care on that measure, it might be skewed by these really long admission patients.

So, I think it is reasonable to do it differently for the kind of safety incidence-based measure versus how your hospital is doing with hyperglycemia management measure.

DR. WINTERSTEIN: That was exactly

our thinking, just to respond to this. And we did have one patient who had three hypoglycemic events in one testing hospital.

CO-CHAIR ROSENZWEIG: Just one comment. I think there had been a suggestion about identifying a blood glucose of less than 40 as being considered to be severe hypoglycemia. And I would caution against that because I think Sue and I were on a committee of the ADA that looked into these definitions.

And the definition for a severe hypoglycemia is generally hypoglycemia, whatever the blood glucose that causes impairment that requires assistance by someone else. And that committee couldn't really come up with a specific number like 40. They came up with a number of 70 as an alert value for hypoglycemia, as I recall.

So, I mean defining it as hypo, I am not against the idea of using 40. I think that is a very good number but to call it

specifically severe hypoglycemia might confuse the issue.

Yes?

MEMBER KIRKMAN: Well, I am concerned, though, about just calling it hypoglycemia because, again, that implies that anything from 41 to 70 is not hypoglycemia.

So, I think it does need to be labeled with some term like severe or really bad or dangerous hypoglycemia.

DR. PACE: How about just having the level in the title, hypoglycemia less than 40?

MEMBER KIRKMAN: Well, I mean I still think -- I don't know. I mean I personally don't mind the term severe because I think it is different for inpatients. And I think the workgroup sort of consider the ambulatory patient, for the most part.

I think severe kind of conveys what we are talking about but I don't know.

MEMBER McCOLLISTER-SLIPP: I would

just be concerned about 70 because I mean what we don't want to do is create some sort of weird adverse incident. Because some people like to hover around 70. I mean they do and God bless them, it is a little uncomfortable but I can understand. I mean there are times when I want to keep it lower than others as well. And I think you need to give some degree of leeway to people who are more comfortable at a lower level than at a higher level. I mean some people prefer to hover around 120 and some people are like just anything that isn't insane.

So, I mean it is pretty easy to go from 80 to 65. I mean, at 65 you are going to do something about it but in terms of dinging a hospital because you go from 80 to 65 or 75 to 65, that would make me a little nervous about what would the adverse effect be in terms of the way you are treated in the hospital.

Because when you are there,

inevitably, the people who are giving you or overseeing your insulin are nurses and they can -- I mean my sisters is a nurse. Nurses are brilliant. I love nurses but some of them can take on a patronizing tone where they try to take your pump away or they -- I mean if the guidelines are too strict, then it creates this sort of negative environment for patients in the hospital, especially those of us who are used to controlling our own glucose.

So, again, 70 makes me a little nervous, 40 makes me a little nervous. I don't know. I mean it sounds like the three bears, just in terms of finding the one that is just right.

But I mean I think we need to think about what does it mean for inpatient experience, especially for those of us who are very neurotic about controlling our blood sugar when we can be, what that is going to mean in terms of patient control.

CO-CHAIR ROSENZWEIG: I think that

is precisely the issue. I mean I think less than 40 is a good number for monitoring for the purposes of this measure. I am not disagreeing with that at all. But people have defined severe hypoglycemia in a variety of different ways and what was clear from that particular committee was that there was no consensus as to what specific blood glucose defined severe hypoglycemia.

Yes?

MEMBER BREEN: I would like to comment on that. Because of that, because there is no consistency, I joke it is like a little bit pregnant. You have just got to call something something at some point.

The concept of the Whipple's Triad was really used in the outpatient community. Meaning, if you are home and you have a patient who goes down and needs assistance, that definition is really developed for ambulatory patients. So I think on the inpatient side, I am very comfortable calling

this severe hypoglycemia because also by definition, they are going to require treatment. Someone with a glucose of 40 is going to get treated by somebody else. So, they fed into that definition of requiring treatment.

So, I think we should just call it severe hypoglycemia.

CO-CHAIR GOLDEN: You want to make a comment?

MEMBER McCOLLISTER-SLIPP: Just again, I mean you get into a question of confusing nomenclature then. Because when it comes to adverse --

MEMBER BREEN: But there is no -- we are saying the nomenclature is already fuzzy. It is fuzzy nomenclature and I think no one -- I doubt that anyone is going to have an issue on the inpatient side calling a glucose of 40 severe hypoglycemia.

CO-CHAIR GOLDEN: I would like to comment that right now we don't have this as

a measurement. This is, again, you get into the issue of perfect versus the good. It is a starting point. It generates data. It generates discussion. It generates a whole raft of, if you will, naval gazing at institutions about the issues we are talking about where people start talking about what you are concerned with.

I don't think we are going to solve this today but it is a place to start. And so, I am comfortable with this as a valid place to start, so, as opposed to sending back.

Yes, Ingrid?

MEMBER DUVA: I was just going to ask Lindsey to explain again the adverse, what you talked about the adverse event. And so can you explain that again? Because then Bill defined adverse event as associated with injury but I thought you were using a looser term.

MS. TIGHE: This is related to the

NQF serious reportable events, which are not measures but actually reportable events by the states.

What I was saying was that the trigger for reporting the hypoglycemia event is patient injury or death, essentially. So, it is that harm factor.

MEMBER McCOLLISTER-SLIPP: And that is what it is for the FDA AERS data, too, which is why it is kind of a joke. Because you can have the same event happen if like it is a pump failure or something like that while you are awake and you catch it at 30 and you are not unconscious but if it happens at night, you become unconscious.

I don't want to confuse the nomenclature by calling something -- I mean I have had lots of 20s before where I was completely conscious and nobody had any idea. I mean I felt horrible and my brain hurt but other than that, that would not, by FDA -- I mean we are not talking about FDA here -- by

FDA standards of that term, it would not be considered a severe hypo, even though on somebody else, they might have been unconscious and could have had a car accident.

CO-CHAIR ROSENZWEIG: Sue.

MEMBER KIRKMAN: So I think I heard the developer say that they are comfortable changing it to severe hypo and that sounds like that is that kind of minor change that we could go with.

So, can we sort of move on? Is anybody that uncomfortable?

CO-CHAIR ROSENZWEIG: I'm not pressing this point. I don't think it is --

MEMBER KIRKMAN: Okay. So I mean is the group consensus changing this and calling it severe hypo is okay for this measure?

DR. PACE: Is there anyone who has an objection to that?

CO-CHAIR ROSENZWEIG: Okay.

DR. PACE: We are talking about

less than 40.

CO-CHAIR ROSENZWEIG: Less than 40.

DR. PACE: We are only talking about less than 40.

CO-CHAIR ROSENZWEIG: Okay. So perhaps we can vote at this point.

MS. BAL: Voting is open.

(Pause.)

MS. BAL: Okay, we have high ten; moderate eight.

CO-CHAIR ROSENZWEIG: Feasibility.

DR. PACE: So, maybe one way to think about this, and we have already talked about feasibility of the other measure is are there any unique issues with this one from the other one, since they are both eMeasures that would be something that needs to be considered?

CO-CHAIR ROSENZWEIG: You know the only unique issue that might be is that you have to collect data on what the patient was

taking, what medications the patient was taking for the previous 12 and 24 hours in relationship to the hypoglycemia.

I don't see any major problems with doing that but that is a different data collection than for the hyperglycemia.

DR. PACE: And the measure developer also did a feasibility assessment for this measure. So, those would be additional data items. And I assume, Kyle, no problem with getting the medication and timing medication in your feasibility assessment?

MR. CAMPBELL: Right, that is correct. This measure scored very similar to the hyperglycemia with an overall average score on our feasibility score card of 2.89.

CO-CHAIR ROSENZWEIG: Okay, any other comments?

MR. CAMPBELL: And I will point out that there are actually fewer data elements here because we are not looking at potential stratification across units. And so

in this case, we don't have to calculate that data element.

CO-CHAIR ROSENZWEIG: Any other comments? Okay, let's vote on feasibility.

MS. BAL: Voting is open.

(Pause.)

MS. BAL: Okay, we have high 15; moderate four.

CO-CHAIR ROSENZWEIG: Usability and use. Any comments by the reviewer?

MEMBER LEDDY: Well, the measure is under consideration for CMS hospital quality reporting program and meaningful use Stage III. No time lines were provided. I would just encourage them to get on with it. We really need it.

(Laughter.)

CO-CHAIR ROSENZWEIG: Any other comments? Okay.

MS. BAL: Voting is open.

(Pause.)

MS. BAL: Okay, high 16; moderate two; insufficient one.

CO-CHAIR ROSENZWEIG: All right, so now we are going to vote on the overall measure. Any other comments before we vote?

DR. PACE: And I would say just to clarify your recommendation of calling this severe and removing the optional, which the developer agreed to, would be what you are voting on.

MS. BAL: Voting is open.

(Pause.)

MS. BAL: Okay, yes, 19.

CO-CHAIR ROSENZWEIG: Okay, very good. Thank you.

CO-CHAIR GOLDEN: I have a question. At some point we do public comments. Do we do that now, do it later?

MS. BAL: Right before lunch.

CO-CHAIR GOLDEN: Right before lunch. So, we would move on to -- I got the wrong day -- so, if it is 11:30, it must be

statins. So, Bob, I think it is yours. Is that right? Okay.

Do we have some folks from CMS who want to talk to us about the adherence measures?

MR. CAMPBELL: Sure, this is Kyle Campbell again from FMQAI.

So, the next three measures submitted for your consideration focus on adherence for patients with diabetes. They were originally endorsed in 2011 as a single measure, with some measures for the three drug classes included, which are statins, ACEI/ARBs and oral diabetes agents.

Recently, NQF recommended to us that the measures are separated but proposed that they remain a pair to keep the reporting and the measures linked. The measures here are different. They are based on administrative claims data, not electronic health record data, the previous measures we have discussed. And as directed by NQF during

the original endorsement period, we worked very closely with the pharmacy quality alliance to establish a standard methodology for NQF-endorsed medication adherence measures.

And the methodology selected, based on extensive testing to establish validity was the proportion of data covered. Therefore, these measures use the proportion of data to cover methodology and are harmonized with the majority of NQF-endorsed adherence measures, including all those that are developed for CMS.

In terms of evidence, it is important to note that the underlying RCTs used to establish the efficacy of these drugs and linking them to improved patient outcomes are relevant and all those studies have protocols to ensure medication adherence.

Regarding the specific selection of the threshold of 80 percent, the majority of studies that link outcomes to adherence do

use this cut point of 80 percent. And once again, this was a decision that was harmonized across chronic medications adherence measures, in the NQF portfolio.

At the direction of the workgroup, we did provide -- we provided supplement to the original evidence form with additional studies. And in each case, there is a clear link between adherence and improved patient outcome.

I think it is also important to note that the denominator for these measures requires at least two prescriptions, which signifies the physician's intent to prescribe and continue the medication and that the adherence is measured across the drug class.

Finally, as clinical practice shifts towards individualized patient goals, rather than standard thresholds across all patients, I think adherence measures are an important tool to assist clinicians in engaging those patients in the management of

their care. Without measures like this, it is very difficult for physicians to determine which patients have filled prescriptions for medications for which they have been prescribed.

And with that, I thank you again for your consideration of this measure. We look forward to any questions you have.

MEMBER BAILEY: It sounds like we have a couple -- how shall we proceed? Answer the questions?

CO-CHAIR GOLDEN: Well we have a number of -- there is lots of questions we could be asking. I think that we will do them sequentially. So, why don't we start with the evidence and just do it sequentially? And the developers will be there.

Because otherwise, I think we could end up talking about all sorts of things that would get into validity and accuracy and all sorts of things. So, let's just do it sequentially and move forward. So, Bob?

MEMBER BAILEY: So, just to confirm, this measure is adherence to statins for individuals for diabetes and this is a process measure using administrative claims. And the goal here is to measure adherence to statins.

CO-CHAIR GOLDEN: Can you move your mike a little closer?

MEMBER BAILEY: So the goal here is to measure adherence to statins using proportionate days covered, which Kyle had already mentioned is now the preferred and harmonized measure of adherence for multiple drug classes.

And so the numerator here is individuals in the denominator with at least two prescriptions of statins with a proportion of days covered of greater than 0.8. And the denominator is patients with diabetes with at least two prescriptions of statins during the measurement period for 12 consecutive months. And so the idea here is to get to patients

that are prescribed statins with the intent of continued therapy.

The exclusions are looking at ways to get a relatively clean population of patients with diabetes, eliminating patients with polycystic ovary disease and steroid-induced diabetes.

So when we look at the evidence, first of all three clinical guidelines are cited. Specifically, the American Diabetes Association, ACE, and the American College of Cardiology. And they specifically mention the role of statins in terms of reducing cardiovascular risk in this patient population. They don't necessarily directly link adherence but there is a tangible link in terms of adherence to statins and the reduction of cardiovascular risk in this patient population.

They also cite several meta-analyses again that point to the efficacy of the reduction of cardiovascular risk in

patients with diabetes that are taking statins. And then they cite several studies that talk about the link of adherence and the reduction in outcomes.

There were some additional studies that were identified by some of the workgroup members that suggested that there was additional evidence creating that link between adherence to statins, not necessarily completely in the diabetic population but with a significant proportion of the population being diabetic. So again, there is a tangible link in terms of adherence to statins and a reduction.

So based on the algorithm in the discussion, we came out with a recommendation that this was moderate evidence to support this measure.

CO-CHAIR GOLDEN: I have a question, again, for the developer. And again, I don't know whether this is evidence or what have you.

You have the evidence about the value of continuation of therapy. If this measure is for accountability, who is accountable and is there evidence about is it a patient function, is it a provider function, is it population specific? I mean, you can put that in a number of categories. And I was just curious if the developer used this as an accountability measure. And if so, is there evidence to support who was accountable?

MR. CAMPBELL: So we have looked at that and we feel that this measure is a shared accountability. So, it would be shared between the patient, the clinician, and then larger organizations, such as the Accountable Care Organizations and plan level.

We do consider this might be, in terms of setting, be considered for the accountable care organization model. They would have responsibility for the overall care of the patient. But we do have the measure specified across plans, Accountable Care

Organizations, and large physician groups.

CO-CHAIR GOLDEN: So that would be helpful. That would put something into usability or something later, or something along those lines.

Okay, Jamie?

CO-CHAIR ROSENZWEIG: Yes, I have a question. It is my understanding that the ADA guidelines for statins include patients from 40 years or older and with an option for patients lower than that, if their LDL cholesterol is less than 100.

So why did you pick 18 years and older as an absolute for all of the patients?

MR. CAMPBELL: Again, in mind with the denominator population already signifying the physician's intent to prescribe the medication. And one of the reasons we select two prescriptions is in case patients had tried a medication and failed it or had some sort of adverse reaction, we take that second refill when we look at the data, the intent to

prescribe over a longer period of time.

And so for that reason, we felt that it was reasonable to consider including all adult patients for which the physician had made a decision that that patient should be on statins.

CO-CHAIR ROSENZWEIG: Well, suppose you have a patient who is 24 years of age and has a very high HDL cholesterol. Why should that patient necessarily be put on a statin?

CO-CHAIR GOLDEN: I think that the measure though is not prescribing it. It is if you are on it, you continue it.

MEMBER BAILEY: Right. So, it is a requirement of two prescriptions within a 12-month period. So, the clinician has already identified that patient warrants therapy and is looking at whether there is adherence to therapy as prescribed.

CO-CHAIR ROSENZWEIG: Oh, okay. Thank you very much.

CO-CHAIR GOLDEN: Are you okay? You are okay down there. Sue.

MEMBER KIRKMAN: So, I sort of thought of it from the other perspective because I think it is clearly looking at adherence when the physician has decided that the patient needed to be on a statin. But it doesn't really address the problem of the guidelines are really that high-risk people should be on statins, regardless of their LDL. But there is still this perception out there that if a diabetic patient who smokes and has hypertension and is 55 has an LDL of 98, they don't need a statin.

So, this isn't really going to get at appropriate prescribing of the statin. It is really just an adherence measure. Is that correct? Because I think that is a little unfortunate.

MR. CAMPBELL: That's correct.

MEMBER BAILEY: I guess this came up in discussion yesterday as well in terms of

looking at the actually laboratory value and I think we determined from NCQA that the sample size is about 411 patients per plan. So, you are looking at small portion of the population in terms of having the laboratory value. Here, you are depending on administrative claims, rather than laboratory values. So, you have a much larger population, which you can draw assumption terms of performance and intervene.

MEMBER KIRKMAN: But it really is just measuring one specific thing, which is adherence, not that the patient should be on a statin.

CO-CHAIR GOLDEN: And I think that would get into importance or impact.

CO-CHAIR ROSENZWEIG: But suppose a person had been prescribed a statin by one physician and then the second physician thought it was not necessary, would that be counted? Would that situation be excluded in this measurement?

CO-CHAIR GOLDEN: That gets into validity. We will get there, too. Okay, good point.

We will ask the developer now. What happens if you stop therapy? How do you count that if somebody has an adverse reaction or another doctor decides this is not an appropriate therapy?

MR. CAMPBELL: And so again, that would be captured by the requirement for two prescriptions. But if the therapy were discontinued, let's say like six or eight months into therapy, that would be a scenario that we would not be able to ascertain from the administrative claims data. And for that reason, they would be picked up in the following measurement year. They would no longer be in the measure in the following measurement year.

CO-CHAIR GOLDEN: Go ahead.

MEMBER McCOLLISTER-SLIPP: Okay, so I will go. So, does that mean that if

somebody starts on a statin and they take it for a couple of months and they start having myalgia as a result of that and they decide they don't want to take it, does that mean -- and you know they may or may not talk about that with their physician. So, the claims data is based on the physician writing a prescription or is it based on the filling of the actual --

CO-CHAIR GOLDEN: It is on the filling. So that patient would actually fall out as a numerator failure.

MEMBER McCOLLISTER-SLIPP: Okay, so that would be a failure, even though the patient had an adverse reaction in their perception that they felt like was enough to make them go off.

CO-CHAIR GOLDEN: As this measure is specified. But we will get to that, yes.

MEMBER DUDL: The comment by Sue is very pertinent. And to set all of these three things in perspective, it does require

two things. It requires adherence is very important. If you are taking a medication because the physician, by criteria and by knowledge, knows it is going to save your life, decrease heart attacks and strokes, then being 20 percent more adherent is like adding a drug that is 20 percent more effective -- another 20 percent. So, it is very important but it is not sufficient.

And of course with the new ACC guidelines, everybody has to rewrite what the criteria will be. And at that point, of course, they have said at 7.5 percent CVD risk type of thing but that will come within a year.

But this needs to go ahead because it can be measured right now and it will be the one important thing that we could continue while we are moving from an LDL of some number to a risk of some number.

CO-CHAIR GOLDEN: So I believe the issue before is if you are on a statin, is

adherence important and is there evidence to support that. I think we will limit it to that.

Are we ready to vote?

MS. BAL: Voting is open.

(Pause.)

MS. BAL: Okay, we have high ten; moderate eight; low one.

CO-CHAIR GOLDEN: So now you get to gaps, an opportunity for improvement.

So, are there variations by race, age, socioeconomics, et cetera in adherence?

MEMBER BAILEY: So, the evidence that was presented by the developer suggests that there is, indeed, a gap and it is based on Medicare data from ten states, using 72 prescription drug plans and slightly more than 7,000 physician groups for 2012. And first of all, the baseline performance at the state level is about 72 percent, the drug plan about 72 percent, and physician groups about 70 percent.

Then when you look at the distribution or the difference between the 10th percentile and the 90th percentile, there is about a 15 percent difference, suggesting that there is indeed a performance gap there.

CO-CHAIR GOLDEN: Okay, anybody want to challenge, question, or go into details on this one? I think we may be ready to vote, unless you had something. No. You have an old card. Okay.

I think we can vote on this one.

MS. BAL: Voting is open.

(Pause.)

MS. BAL: Okay, we have high 15; moderate four.

CO-CHAIR GOLDEN: So, priority. And I guess the question before us, it is an interesting question, how you define what we are going to vote on here is adherence a priority item for impact in the management of this issue. Is that how to frame it?

MEMBER BAILEY: I guess I would

frame it in a slightly different perspective and it is based on the evidence.

So, given the high burden of diabetes in the U.S. population and the high burden of cardiovascular disease and the very tangible link between adherence to statins and the reduction in cardiovascular events in this patient population, I would advocate that this is a high priority condition.

CO-CHAIR GOLDEN: Okay.

MEMBER TAYLOR: I don't think there is any question that you are not going to get the benefit of statins if you don't take them. There is also no question that if you have had coronary disease, you benefit from being on a statin.

People with diabetes with complications have been shown in cards to benefit from being on a statin. People with diabetes and that all had LLT, failed to show any benefit in primary prevention.

And the American Diabetes

Association -- I just went online to look up, you guys probably wrote this stuff so you can correct me, but the American Diabetes Association says in response to the new American Heart Association, American College of Cardiology recommendations, which are mired in the difficulty of the inaccurate risk calculator that they put out, that American Diabetes Association is reconsidering their position and especially is not taking a position on whether statins are indicated for all people in primary prevention between the ages of 40 and 75.

So for us to come out and say it is important to take your statin, embedded in that might be taking statins for people where at least some of the experts are telling us they are not sure that statins are needed.

CO-CHAIR GOLDEN: Sue.

MEMBER KIRKMAN: So, right before the standards of care went to press, that is when the AHA/ACC formerly NIH guidelines came

out. And so I think that statement was just put in there to say just because people would say well, why didn't you reference that or whatever. I don't think it is true that the ADA is reconsidering its position because actually the ADA position for quite a number of years has been very similar, which is most high-risk people with diabetes should be on a statin, regardless of what their LDL is.

So when the AHA guidelines came out, to me it wasn't really any different for people with diabetes. So, the recommendation of the ADA is if you are over 40 and have one other cardiovascular risk factor, which the vast majority of people with diabetes, particularly Type 2 diabetes have, you should be on a statin, regardless of your LDL.

So I don't think -- and the primary prevention data is fairly strong, if you look at the meta-analyses of all of the statin trials that were either in people with diabetes or looked at diabetes as a subset,

there is pretty strong data for cardiovascular risk reduction. I think it is not as strong for mortality reduction as the secondary prevention, if you have already had a heart attack, but it is pretty strong for primary prevention in high risk diabetics. So, I don't think it is the case that there is really not good evidence or that this is a change from what ADA has recommended.

MEMBER TAYLOR: Could I read what is on the ADA website? It is just one sentence. It says -- this is about the new guidelines from the American College of Cardiology, American Heart Association, that had been the NHLBI. The Association will consider whether moderate dose statins should be used for the primary prevention in all patients 40 to 75 years of age with diabetes, regardless of baseline lipid levels or the presence of other cardio risk factors, notably the revised guidelines and so on. And they go on to talk about --

CO-CHAIR GOLDEN: And maybe the key word there is moderate because they might change it to high dose.

MEMBER KIRKMAN: Right but the current recommendations are very consistent. It is just I don't think it uses the term moderate. And I think the person, it was a staff person, that wrote that little paragraph, I don't think she meant that AHA guidelines are really different and we are going to reconsider the position.

So, I think it is pretty consistent, except that the ADA guidelines did not say moderate dose statin.

MEMBER McCOLLISTER-SLIPP: Yes, and I think it is really helpful to have Sue giving us the context of how that got in there. It one thing to read it without a context but I think within this context, it makes perfect sense.

CO-CHAIR GOLDEN: Are we ready to vote?

MEMBER KIRKMAN: I personally, wouldn't have put it in there, even.

CO-CHAIR GOLDEN: Are we ready to vote? Yes.

MS. BAL: Voting is open.

(Pause.)

MS. BAL: Okay, we have high 14; moderate four; low one.

CO-CHAIR GOLDEN: Okay, now we get into the fun stuff. So now we get into reliability of the specifications and reliability testing.

So, let's make sure we are clear. This is going to be if the data are collected, do you actually collect it accurately as opposed to the validity, which we will get into some of the other issues as to what the data means. Correct?

So, Bob, did you want to -- is that a good context?

MEMBER BAILEY: Sure. And so

empiric reliability testing was performed by the developer. And the reliability testing was done at three different levels, the state level, the prescription -- actually four levels. The prescription drug plan level, the physician group level, and the accountable care organization level. And the minimum threshold for reliability was met for all of these different groups of analysis.

CO-CHAIR GOLDEN: So when collected, it was collected accurately.

Comments or questions? Shall we vote?

MS. BAL: Voting is open.

CO-CHAIR GOLDEN: We have one listing.

MS. BAL: Okay, thank you.

(Pause.)

MS. BAL: Okay, the vote is high 14; moderate four.

CO-CHAIR GOLDEN: So now we get to validity.

MEMBER BAILEY: So the validity testing was done by a systematic assessment of face validity, where they convene a technical expert panel. And 77 or 78 percent of the panel either agreed or strongly agreed with the validity of the measures.

CO-CHAIR GOLDEN: Okay. So, let me just take a look here. All right, Bill, we will start with you and we can talk from there.

MEMBER CURRY: The one area that I have a question about is missing data. And there are patients who I could write a prescription to start a statin and they could get a fill of that or a couple fills of that at their local pharmacy and then they go off to the VA to get their care. They could buy it as a low-cost option at a large box store. They could go to the grocery store that is offering atorvastatin for free to get people in the door. And now we are not going to capture that they are going to continue to

have these prescriptions being filled. And so they might have initial two prescriptions but the coverage time might be less --

CO-CHAIR GOLDEN: Let me stop you there for a second because that is a problem with reliability. We may have to reconsider reliability. In fact, I am going to consider that we do so because what you have just brought up is that if people pay cash, there is no claim.

MEMBER BAILEY: Good point. And if I could just raise here that the developer did present the information addressing two major aspects that were threats to validity and the reliability, specifically the cash prescriptions and that it was minimal impact on the data in the analysis. And they also looked at missing availability of the day's supply. So, two of the major threats.

They did not address the one, specifically the patient going to the VA and getting a prescription for a portion of the

time as well.

CO-CHAIR GOLDEN: And the other question I had, so I think I talked about this earlier. So, is it the total days' supply or the total number of pills that is in the measure?

I am asking for the developer, when you do the collection, is it --

MR. CAMPBELL: Yes, so we looked at both. We looked at the first prescription within the measurement period to the last prescription in the measurement period plus the days supplied. Then it looks at the number of days that are actually covered by the medication, if you added up all those days supplied across all the prescriptions.

CO-CHAIR GOLDEN: So my question for you is if you prescribe something every other day, would that fail or would that be acceptable?

MR. CAMPBELL: It would be acceptable, as long as the every other day was

indicated, in terms of the days supplied. And in most cases, I believe that would be correct, unless the patient were prescribed it one way and the measure was still -- and it was still one way versus what they were actually told to by the physician. But if they are following the instruction, it should be captured.

CO-CHAIR GOLDEN: Jessie, do you want to talk about reliability or talk about validity?

MEMBER SULLIVAN: Well, this is reliability.

CO-CHAIR GOLDEN: Okay.

MEMBER SULLIVAN: So I am concerned about the cash and I don't understand, if the developer could explain, how you said that wasn't a problem. How did you determine --

MEMBER BAILEY: I should just clarify, too, that that falls under the threats to validity section. So, this is the

appropriate --

MEMBER SULLIVAN: Oh, this is the validity.

Okay, so I have several questions about validity. One is the age. It seems to me that we are saying that it is a good thing that we have this measure. So, I don't understand why it should start at 18 instead of at 40, from what I have been hearing people say. So, that concerns me.

I am concerned that the level of accountability for this measure is at the physician level but, as I understand the measure, it doesn't take into account new starts. So, if you start someone on the medication, a new start in November, you are going to get dinged because they wouldn't have been on it for 80 percent of the measurement year. And if they have an adverse reaction, then you stop it because they had an adverse reaction. That is going to count as a ding.

And if you are measuring at the

whole population level of the whole country, that doesn't matter because it is trivial. But if you are measuring in one practitioner's practice, it can matter.

And then I am concerned about patients with no pharmacy coverage who again, if you are looking at the effect for patients, it doesn't matter. Then we know that patients don't have pharmacy coverage. But if you are holding the physician accountable, that is going to be an issue.

And then I particularly don't understand about the missing data because I know that Walmart doesn't contribute to Surescripts. So, you are not going to get fill data -- I mean if you go to CVS and pay cash, it is still going to be in the Surescripts data. But if you go to Walmart and you pay cash, it is not. Right now that is the case. So my patients at the community health centers they -- those are all, I think, threats to validity.

CO-CHAIR GOLDEN: We are having some background noise on the phone, too. So, please be careful with paper and what have you.

So, the developers, can you respond to some of those issues?

MR. CAMPBELL: Yes, absolutely. So with regard to the age issue, I think again it is important to note that in the denominator the physician has prescribed the medication and we do have evidence of a second fill. And so for that reason, we would suggest that this would be applicable to all adult-aged patients greater than 18.

In terms of the level of accountability, it is not at the individual physician level but at the physician group level. And in terms of new starts, if a patient were, let's say they started in October and they had fills in October, November, and December, if they were adherent, you would capture that because the follow-up

starts with the index event, which is their first prescription. So, you will not be dinged in the measure if the patient doesn't start the medication until late.

In terms of the adverse reactions, so again I think it was brought up the potential for side effects with statins being myalgia. While we do know that occurs, we also know that the goal would be to try and keep patients on statins. And so the switching of patient from one statin to another also is captured. So, you get credit for if you start on let's say Lipitor and you switch to Lovastatin, as long as the patient is continually adherent, those count in the total proportion of days covered and the prescriptions are adjusted accordingly, according to the algorithm.

And then as far as the cash prescriptions go, we have done a limited sensitivity analysis with regard to this, where we tried to simulate what would happen

if patients didn't have claims for medications on the discount formulary, what would happen to measure rates. And we didn't see any differences but we didn't have the ability to test it with external data. However, I will point out again that you are comparing yourself to other providers' mean scores. And so, therefore, if there are any issues with cash prescriptions, you know those will be similar across the board.

And as mentioned, we do know that there is an effort to fill some of the gap with the cash prescription information. But it is not 100 percent and it won't be -- you can probably never expect that it will be 100 percent perfect in terms of a measure but I think it is the best we have and it certainly, given the performance gap, ample room for improvement that is not related to the use of cash prescription.

MR. MATTKE: And one more comment, Soeren Mattke from RAND, for the developers.

In order for the cash prescriptions to have a big effect, you would have to assume that a patient goes back and forth between a claims recorded prescription and a cash prescription because if they only went to Walmart and got the four lovastatin, we wouldn't pick them up in the denominators. So we would really only pick them up if they got two prescriptions through their benefit and then rest of them through Walmart and then potentially went back to the pharmacy in the subsequent years.

So, we don't expect this to be a large impact given how patients usually behave, which is they fill all the prescriptions in the same place and repeatedly.

MEMBER SULLIVAN: I just want to say that that is not true in the Medicaid population.

So, people gain coverage, lose coverage, they stay a patient of the doctor,

they go in and out of coverage on a plan. And those patients are not randomly distributed among doctors in the U.S. They are definitely non-randomly distributed.

MR. MATTKE: No, but those would fall, if they lose coverage, they fall out of our denominator file.

MR. CAMPBELL: Yes, that is correct. Just to clarify that point, this measure requires 11 of 12 months continuous eligibility. So, they have to be continuously eligible for Part D, Medicare Part D, for them to be considered in the denominator.

DR. PACE: How could it go to age 18 if it is only people on Medicare?

MR. CAMPBELL: Well there are dual eligibles with Medicare. So, patients that are not part of Medicare that are aged in. So, this would include those dual eligible Medicare/Medicaid patients.

CO-CHAIR GOLDEN: So, let's make it clear. This is a Medicare-only measure?

MR. CAMPBELL: It would include Medicare and Medicaid for Medicaid patients that are dual eligible for both.

DR. PACE: Those patients enrolled in Medicare Part D is what you said. Correct?

MR. CAMPBELL: That is correct. Those patients that have continuous enrollment in Part D that would have no more than a one-month gap in coverage.

CO-CHAIR ROSENZWEIG: I just would have an issue again with the specifications. First of all, if you have patients, and as far as I understand it if you have dual eligibles, you will have patients who are under Medicare who are younger than age 40, first of all the measure indicates only a denominator exclusion for gestational diabetes but does not specify patients with preexisting diabetes who become pregnant, which I think needs to be changed or needs to be added.

Clearly, a woman who has preexisting diabetes should not be on statins

if she becomes pregnant. Obviously, if they are taken off of statins for pregnancy, it would be considered to be nonadherence by the way this measure is set up.

So, I think that is a potential problem. The second problem is in the under age 40 population, half of the patients that you are dealing with in that population are patients who are women. And I think this measure might inadvertently increase the risk of women with diabetes being on statins during pregnancy. And I think that is a potential problem with the measure as it is currently constructed.

So, I would actually urge you to actually move the age up to 40 for that basis.

CO-CHAIR GOLDEN: I guess the last issue would be something with usability and adverse consequences.

CO-CHAIR ROSENZWEIG: Correct but we are talking about specifications here.

MEMBER BAILEY: And the other

consideration here is that the clinician and the patient have already had the discussion that the therapy is appropriate and therapy has been initiated and we are looking at continuation, not whether it is initiated.

DR. PACE: You have to keep in mind the denominator are people that the physician has already prescribed this, based on guidelines, et cetera, in discussion with the patient.

CO-CHAIR ROSENZWEIG: Correct but the patient might become pregnant during the period of time afterwards.

CO-CHAIR GOLDEN: Do we have other comments or questions or are people almost ready to vote?

MR. CAMPBELL: This is Kyle Campbell again from FMQAI. Just a comment to the pregnancy issue. This was something that we did look into as part of the testing and just the part is part of our ten-state data sample for this population, the occurrence of

pregnancy, at least according to the administrative claims data, was exceedingly rare. And so for that reason, we did not develop an exclusion specifically for pregnancy.

CO-CHAIR GOLDEN: Jessie?

MEMBER SULLIVAN: Well, I guess I did see the Part D thing, but it hadn't clicked in my mind. And I think Helen is right, someone who is dual eligible isn't going to be on Part D. They are going to have their drug coverage through Medicaid. So, really we are just talking about the Medicare population but it doesn't really say that.

And I guess my concern is that if this gets endorsed by the NQF and it is hanging out there, even though in the specs it is intended just for the Medicare population, it doesn't read that way. And I think that there are risks in this, in the younger age group that we have identified. I think it is probably a good Medicare measure and I am not

sure that it is a good Medicaid measure for young women in their 20s. I think it could do harm.

DR. PACE: But the specifications specifically say continuously enrolled in Part D. So, that is part of the specifications.

MEMBER SULLIVAN: I know but I think people won't necessarily read that. When I am looking for a measure, if I am thinking I want to do something about adherence to statins and this looks like to me the measure and that is not my population, I might start using this measure. I understand but people don't -- even the NCQA people yesterday, they said we know the measures get used in different ways for different populations. And they do.

CO-CHAIR GOLDEN: Sue.

MEMBER KIRKMAN: But I don't think this measure is going to push more people onto statins. So, I don't really understand the concern that this is going to make more child-

bearing potential women go on to statins. It is really if you have decided that your patient should be on a statin, whether they are taking it or not.

So, am I missing something there?

MEMBER McDERMOTT: No. I was going to say I think it is the women that become pregnant, they have their two doses and then they become pregnant. Now they are off and they are going to look noncompliant because that is a good thing for them to be off.

MEMBER KIRKMAN: But again, I mean --

MEMBER McDERMOTT: And they are still continuously --

MEMBER KIRKMAN: -- if you talk about the total universe of Medicare Part D patients, that is going to be one in 50,000 people. And so, I mean again, you can't make an exclusion for every single thing, especially if it is really rare.

CO-CHAIR GOLDEN: So we might want to change the title to adherence to statins for Medicaid individuals with diabetes or Medicaid recipients.

MEMBER KIRKMAN: Medicare or Medicaid?

CO-CHAIR GOLDEN: Medicare.

MEMBER KIRKMAN: But I am concerned that other people may adopt this, other health plans and so forth. But I don't think it is going to make more people go on statins, necessarily. I think it is just going to make people adhere better.

CO-CHAIR GOLDEN: I am getting people antsy to vote. So, I am seeing people who want to vote.

So, shall we vote? Let's vote.

MS. BAL: Voting is open.

CO-CHAIR GOLDEN: This will be an interesting vote.

(Pause.)

MS. BAL: Okay, we have high

three; moderate 13; low one; insufficient, two. And that passes.

CO-CHAIR GOLDEN: This is -- what is this one here? Composite. Feasibility.

MEMBER BAILEY: So feasibility, the data that is required here comes from administrative claims data bases. So, with a few exceptions that we already talked about, specifically cash claims, that this data is routinely generated during the course of care because everyone wants to get paid and there is no significant burden in terms of collecting this information.

CO-CHAIR GOLDEN: -- a replay of our last conversation but go ahead.

MEMBER TAYLOR: Just a point of clarification. Our process is, if we did want to suggest the name change that said for Medicare-eligible patients or something, so it is prominently displayed because that might change the way some of us vote about these, we ask the developer if they would find that

acceptable. Is that the process?

CO-CHAIR GOLDEN: That would be fine. Did you have any objections to that, changing the name? I will ask the developers. I mean, it is not a specification change it is just that it is a label.

MR. CAMPBELL: We can definitely consider that in consultation with CMS.

CO-CHAIR GOLDEN: And CMS is not here?

MS. RICKSECKER: This is the CMS GTL for the contract. I would be agreeable to changing the name of the measure to accommodate that and including Medicare-eligible individuals in the title.

CO-CHAIR GOLDEN: Great. Thank you.

So, we are now feasibility. Any other comments or do we want to vote? Okay.

MS. BAL: Voting is open.

(Pause.)

MS. BAL: Okay, we have high 14;

moderate four; low one.

CO-CHAIR GOLDEN: Usability and issues of is the measure publicly reported. This is I think where you get into adverse consequences or unintended consequences.

So, currently this measure is not in use. The goal is for public reporting of the measure. And we have already talked about some of the concerns about potential unintended consequences but no other ones were identified.

MEMBER KIRKMAN: So, I guess somebody brought this up earlier and there was some discussion yesterday, too, but I think if you are talking about clinician reporting, clinician-level reporting, I mean it is just a concern that there is only -- there are things that you can do to increase adherence but Jessie may have a population of patients that are just inherently going to be less adherent than some Beverly Hills private practice person.

So, I mean it is -- and again, I think it kind of speaks to the quality of our whole health system more than individual level quality of care -- individual clinician-level quality of care.

CO-CHAIR GOLDEN: To follow-up on your comment, we have been going out dealing with total cost of care and episodes of care in Arkansas and I have been doing Town Halls. And the first set of hands that go up will be what about the patients? And what about -- they are component to this and why am I being held accountable?

And we talked about a number of ways to mitigate that but I can, to follow-up on that, I would say that if we are not careful with this measure and it is reported in a way that clinicians are held accountable, it will result in adverse patient selection because the docs will hate it and they will start making changes. Patients already in my neighborhood do get fired, if you will, for

non-adherence and lawyers will tell them get rid of folks for non-adherence.

So, if it is framed wrong or if it is, in my community at least, there will be a number of clinicians that will avoid patients.

MEMBER DUDL: Low adherence, however, is workable. And if you don't measure it and you don't go after it, it doesn't improve. Low adherence can be helped by looking at everybody that you treat because you are looking at the percentage and finding those that are actually easily moved but didn't come in. It can be worked by learning how to handle patient barriers and there are four or five simple things. And we moved adherence up from 50 percent to 85 percent in three years. If you focus on it, you can do it. I don't think there is that much adverse selection. We did not find that much adverse selection in our group. I'm sorry. We didn't find people leaving. Nobody ever gets booted out.

So, I think those are overly worrisome things when you have a really big potential for adhering.

MEMBER MAKAROFF: I just want to echo what has been said about the adverse selection. I think like in a place where you have a pharmacy across the street or in the building, it is going to probably incentivize or those plans are going to look better.

And like a health center population, Medicaid population, where I see patients it is a bus ride to the pharmacy. My patients are not adherent for a lot of reasons. And a lot of things we do. I have a case manager. I have a social worker. I give them a bus token. But I don't know how that is going to affect our performance as the health center, who is really working hard. I don't know if it is going to capture that quality of care.

CO-CHAIR GOLDEN: To follow-up on my comments well, not that you can't fix it,

it is a matter of how the measure is used. So, I have seen useful measures used in a punitive way. That is my concern.

MEMBER McCOLLISTER-SLIPP: And that is sort of my question maybe for NQF on that front is, I mean how -- I know that we can't really predict how these measures are going to be used. They will be used for some things that we now about but others that we won't.

I mean something like this is absolutely something we should shoot for from a compliance perspective, from a health system. But I think this makes perfect sense for like an Accountable Care Organization but for an individual practitioner that is not part of an Accountable Care Organization, that might be slightly different.

DR. PACE: So, let's clarify. I think Kyle said it is not for an individual clinician. What is the smallest unit of analysis?

MEMBER BAILEY: Physician groups.

MR. CAMPBELL: That is correct, physician groups. And it would mostly likely be limited to have a certain patient population size within those groups.

MEMBER BAILEY: So a quick question for you, Kyle. Was this first intended use for the physician value-based modifier program where it is physician groups above a certain size?

MR. CAMPBELL: CMS hasn't made any decisions with regard to its use at this point. It was submitted in the measures under consideration list at this point for the ACO Shared Savings Program.

CO-CHAIR GOLDEN: Yes, actually I think, and correct me if I am wrong, Kyle, but I believe there is an adherence measure in the CPCPI program that you have to -- that is used in determining whether or not you get shared savings.

MR. CAMPBELL: I'm not aware of an

adherence measure currently in the CMS ACO Shared Savings Program. I can tell you that our other measures are in use. This measure NQF 0543 that looks at the coronary artery disease population and measures adherence to statins and is entirely harmonized with this one is currently in use by the quality resource use reporting program, which provides individualized report to physician concerning their patient population.

CO-CHAIR GOLDEN: Yes, I am pretty certain they were all financial implications with some of these measures.

Tracy?

MEMBER BREEN: I just want to say this always comes down to this risk-benefit ratio and I think there are potential risks of patient dumping that are just going to be out there. But when you look at the benefits, if you don't start measuring something, you can't make changes. And in fact, the places have demonstrated that they can move, potentially,

adherence rates and we have already assessed that adherence to these medications results in better outcomes and lives saved. That is the weight but I have the same concerns about the avoidance of challenging patients. Although, again, this measure Medicare-eligible patients. We keep going back to who this is specifically looking at, a slightly different group of patients.

MEMBER BAILEY: And if I could just expand upon that, so it is physician groups that will be compared with physician groups, ACOs compared with ACOs. So, you have comparison with comparable groups and so you have an idea of where you are performing with respect to your peers and identify opportunity because we all think we do well at everything by nature of being physicians or healthcare providers. But when you start to understand what you are doing compared to your peers, it incents behavior changes to improve outcomes.

MEMBER KIRKMAN: So not to belabor

this too much but I agree for quality improvement, this is an incredibly important measure but again, we are talking about public reporting here. And even saying your peers, so I work in an academic clinic at a state university. We take all comers. I can tell you the private practice endocrinologists are very good at knowing that they can get rid of people and that UNC won't refuse anyone.

So I just am concerned a little bit about a publicly reported measure where some things are within the physicians or the group's control but a lot may be the patient population. And I guess we will just have to see what happens.

MEMBER BAILEY: And to address that concern, it wasn't included in the evidence provided but there was an analysis that was presented at Academy Health a few years ago that looked at whether changes in performance against certain quality measures, and the focus was on diabetes, whether it was

patient characteristic. So they controlled for various patient aspects. And then they did it at the physician level as well. And performance actually tracked at the physician level and not at the patient characteristic level, suggesting that there are behavior changes that occur when something is being measured.

MS. BAL: Voting is open.

(Pause.)

MS. BAL: Okay, we have high five; moderate ten; low four. So, it does go through.

CO-CHAIR GOLDEN: And a global picture.

CO-CHAIR ROSENZWEIG: I just want to make one comment before we vote on this. And that is I don't quite understand why if gestational diabetes is listed as a denominator exclusion, why diabetes and pregnancy can't also be added as a safety factor.

It affects the same population and may involve very small numbers of people but I think it is a reasonable thing to suggest that the developers add that one additional exclusion.

MR. CAMPBELL: This is Kyle Campbell from FMQAI. Just the genesis of the gestational diabetes was to harmonize the diabetes identification algorithm with NCQA, which has a number of diabetes measures.

But if the Steering Committee feels strongly that pregnancy should be added as a safety concern, we would be amenable to adding that exclusion to the measure specification.

CO-CHAIR GOLDEN: Yes?

MEMBER KIRKMAN: So again, I think GDM is excluded not because it is bad to give women with gestational diabetes statins but it is because the denominator is sort of everybody with diabetes. And that is sort of not considered diabetes for a lot of these

measures.

So again, for pregnancy, I mean the numbers are going to be so low, a Medicare Part D patient with preexisting diabetes, that gets a statin -- you know that gets a prescription for a statin inappropriately, that it is not -- I mean my understanding it is not worth programming in exclusions for incredibly rare events.

DR. PACE: And that is actually part of the NQF guidance. And we ask developers to really even when they do specify exclusions, to do some analysis to see whether it is sufficient enough to warrant data collection and programming to do that.

And so I think the phrase that Helen uses is decimal dust. So, we need to think about whether that, you know first of all, is that going to be kind of a random event and as you said incredibly small versus something that is really going to skew performance results.

CO-CHAIR ROSENZWEIG: I'm not sure this is decimal dust, frankly. And I am a little concerned about the fact that the measure might inadvertently be applied to people outside of Medicare Part D.

MEMBER McDERMOTT: I was going to say from someone that uses measures a lot and we need to use external standards, this is the kind of measure that one would like to use in health plans. And we have ways of -- we have adherence type measures today within our building but they are not coming from an external entity.

The thing that is wonderful is that this has been standardized and harmonized and so forth and so on. So, this is the kind of thing that those of us that need to be using standardized measures would love to be able to grab onto.

So to me, to make the label that is only for Medicare Part D is kind of disappointing but that doesn't mean we

couldn't use it.

CO-CHAIR GOLDEN: So, I am going to try to corral us here. We have just gone through a series of votes. And a lot of the issues being raised here were incorporated in those series of votes on usability and reliability.

So, my suggestion here is this is a yes or a no. Most of the folks in the room have been saying things are okay. I mean you are free to vote yes or no. If we vote it down, we might revisit some things to find out okay what was the no all about. But I think continuing this conversation at this stage is not going to result in any significant change.

So unless using my prerogative -- I see people raising their hands to vote. But unless people want to object, I would suggest we vote on the overall measure.

MS. BAL: Voting is open.

(Pause.)

MS. BAL: Okay, we have yes, 15;

no, four. The measure passes.

CO-CHAIR GOLDEN: Okay. It is about lunch time but we have public commentary. So, we are going to let the non-committee members discuss anything of their choosing.

MS. TIGHE: Operator, if you could see if anyone on the line has a comment and anyone in the -- no one in the room.

OPERATOR: At this time, if you have a question or a comment, please * then the number one on your telephone keypad.

And there are no questions or comments at this time.

CO-CHAIR GOLDEN: Okay. We are inching our way forward. I am going to suggest that we take ten or 15 minutes but eat during the next discussions, so we keep things moving.

So, is lunch here?

MS. TIGHE: It is.

CO-CHAIR GOLDEN: All right. So,

folks can gather about and let's say about quarter of, we will reconvene.

(Whereupon, the above-entitled matter went off the record at 12:31 p.m. and resumed at 12:48 p.m.)

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

(12:48 p.m.)

CO-CHAIR GOLDEN: So, if we get people's attention, we will move along.

So, we are talking about adherence to what? Adherence to a blue box.

MEMBER DUDL: ACEI/ARBs.

CO-CHAIR GOLDEN: ACEI and ARBs, okay. I think that we heard the developer talk in general about these. Does the developer want to make specific comments related to ACEIs and ARBs or shall we just go straight to discussing the evidence?

MR. CAMPBELL: I think it would be fine to go straight to discussing the evidence. These measures are structured in exactly the same way and harmonized.

MEMBER DUDL: These are Siamese twins joined at the hip. So, thank you, Bob, for doing all that heavy lifting. This should go fairly simply.

Just quickly, this is a process

measure which takes us from Box 1 to 3, 80 percent adherence to ACEI/ARBs in people with diabetes where the use of ACEIs has already been shown and systematic reviews, three of them, but there is at least moderate evidence of decreased CVD events with their use.

Clearly, if the issue is just should we measure it, the evidence is high or moderate. However, the proposal is way to measure and it is the 80 percent adherence as target, which is a distal step, which takes us down to box 7.

There were three articles then submitted or there are three articles that suggest adherence makes a CVD difference. Two submitted by the submitter, which adds empirical evidence. There was an AJMC article a little bit older showing that high adherence at 67 percent produced at 23 per thousand versus a low adherence of 15 per thousand difference at hospital MIs and strokes after three years. So, that was good circumstantial

evidence. Not the 80 percent but there were two other articles at 80 percent that I did find that do speak to this. And these are mixtures of diabetics and CVD because they were all looking at higher CVD risk people.

And the first was very good. It was an Aetna study with Harvard Brigham, the MI FREEE trial, where they did secondary analysis and they looked at 80 percent of proportion of days covered, exactly what we are looking at with ACEI/ARBs and showed a 24 percent difference from partial adherence, 40 to 80 percent, which was no bigger than placebo. So, this would suggest the 80 percent was a really important cut point.

A second study -- however, they were not all diabetic so we had to move on to a second study, which was all diabetics with heart disease by Ho in BMC Cardiovascular Disorders, 4,000 patient retrospective cohort studies with ACEIs and statins. The greater than 80 percent adherence, same proportion of

days covered had a 45 percent drop in all-cause mortality compared to low adherence. But here they had a proportionate drop down to 50 percent, rather than the sudden drop off.

However, there were no studies that were submitted that said that it wasn't important and the group, I believe, thought that this worked through to rating this -- oh, and the other thing, of course, is fairly high certainty of clear benefit versus non in box 9, gives us a moderate rating.

And I wonder if Anne or anybody else who is on the call want to speak about that.

CO-CHAIR GOLDEN: I have a question for you. So, you have some data that shows adherence is associated with some better outcomes. But could that be because you have blood pressure control, so you are just measuring something twice?

MEMBER DUDL: That could be. In fact, probably is. They probably do work

through blood pressure control but that is not the question. It is actually, in a way, easier to measure adherence in some ways than it is to try to get repeat blood pressures, if we have the data.

So, I think the question in front of us for adherence is a very good one. Clearly, as you point out, maybe they would work better with DASH diet or do some other lifestyle things that would drop blood pressure and make a difference. But these three studies really looked at it this way and their evidence was pretty good.

CO-CHAIR GOLDEN: Everybody is eating or they are ready to vote.

MEMBER DUDL: This is the perfect time to talk.

CO-CHAIR GOLDEN: Ready to vote?

MS. BAL: Voting is open. You have 38 seconds to vote.

CO-CHAIR GOLDEN: Are we missing anybody? Is everybody here?

(Pause.)

MS. BAL: Okay, we have high six; moderate 12.

MEMBER DUDL: Are we at gaps in care? They clearly demonstrated gaps in care as was presented earlier. The same data.

And there was a disparity gap in care with Hispanics.

The priority, again, this would be high prevalence, high severity, high cost, high priority.

MS. BAL: So, are we ready to vote? Okay, voting is open.

(Pause.)

MS. BAL: Okay, we have high 15; moderate three; low one.

MEMBER DUDL: And then we go on to reliability.

CO-CHAIR GOLDEN: No, I think we are doing priority.

MEMBER DUDL: Oh, I'm sorry, priority.

Okay, I covered priority. It was high cost, high severity, high prevalence, high priority.

MS. BAL: So, is everybody ready to vote?

MEMBER DUDL: I'm sorry.

MS. BAL: Is everybody ready to vote? Voting is open.

(Pause.)

MS. BAL: Okay, we have high 14; moderate three.

MEMBER DUDL: Reliability, the same studies were presented, I think, earlier. The reliability test 0.82 for states is very high; physician groups, 0.74; drug plans, 0.76. So, they were all of at least moderate reliability.

MEMBER SULLIVAN: So, I would be more comfortable with the precision of the specifications, if we stated that the title of the measure would also include Medicare, as for the previous one. I do understand it is

in the details but I think it calls it out.

CO-CHAIR GOLDEN: Yes, so we have a series of measures from CMS. So the request to CMS would be to change all of them to a Medicare measure. Okay, thank you. I would assume that would be reasonable, if they have already agreed to the first one.

MS. RICKSECKER: This is Elizabeth Ricksecker, the GTO from CMS working on this contract with FMQAI. And CMS would certainly be agreeable to adding Medicare-eligible to the title of all of these measures. Thank you.

CO-CHAIR GOLDEN: Thank you.

MEMBER DUDL: Good point, thanks.

MS. BAL: Voting is open.

(Pause.)

MS. BAL: Okay, we have high ten; moderate nine.

MEMBER DUDL: We are at validity. Face validity was strong, 77.8 percent

together exclusions were as noted. Scientific acceptability was clear. No problems with code definitions or specifications. Elements were clearly defined. Logic calculation is good and could likely be implemented.

I think we found that it would be at last moderate.

CO-CHAIR GOLDEN: Comments, questions? Open the pools.

MS. BAL: Voting is open.

(Pause.)

MS. BAL: Okay, we have high five; moderate 13; low one.

MEMBER DUDL: Next, I think is feasibility. Pharmacy claims are used and it is deemed high feasibility.

CO-CHAIR GOLDEN: Ready to vote?

MS. BAL: Polls are open.

(Pause.)

CO-CHAIR GOLDEN: So, I don't know, Jim. I think that we either have to have you do all the measures or we have to be

fed earlier.

(Laughter.)

MEMBER DUDL: And I have to come after Bob every time. He had three years to come back.

MS. BAL: Okay, we have high 16; moderate three.

MEMBER KIRKMAN: So, I think part of it is that the measures are so similar, this one and the last one and the next one. The issues are really the same that we have already discussed, but it is also that Jim is wonderful.

(Laughter.)

MEMBER DUDL: Thank you but I think that it is that they are the same. The evidence is there.

CO-CHAIR ROSENZWEIG: The one issue here I think -- I don't think it makes a difference in terms of the measure but there are more options to ACEI/ARBs with respect to treatment than there are with respect to the

statins. I mean, you --

MEMBER DUDL: Well, do you mean between many different drugs or do you mean other anti-hypertensive agents, like hydrochlorothiazide?

CO-CHAIR ROSENZWEIG: Agents, yes.

MEMBER DUDL: Hydrochlorothiazide and beta blockers -- well not beta blockers too much, but beta blockers actually with CVD and calcium channels. They all drop CVD a lot. So, you know I think there is a good point there.

CO-CHAIR ROSENZWEIG: And the evidence of the ACEI and ARBs in African Americans is less solid than it is in the other populations --

MEMBER DUDL: Correct.

CO-CHAIR ROSENZWEIG: -- for preferential drugs.

MEMBER DUDL: That is correct.

MEMBER KIRKMAN: But again, this isn't a measure to move more people to these

drugs. This is a measure to say that if you thought your patient should be on the drug, are they adhering?

CO-CHAIR ROSENZWEIG: Correct. I fully understand that.

MEMBER DUDL: That is correct. And so the drugs have been shown to be very powerful when used as indicated and adherence seems to be important in that group. And that is what we are after.

Usability, I think that, again, as planned use and not in-use, but I think the discussion last time carried the data. It is at least moderately usable.

CO-CHAIR GOLDEN: Other comments? All right.

MS. BAL: The polls are open.

(Pause.)

MS. BAL: High eight; moderate ten; low one.

Do we want to vote or talk?

CO-CHAIR GOLDEN: Time to vote.

(Laughter.)

MS. BAL: Okay, polls are open.

(Pause.)

MS. BAL: We have yes 18 and no one, passing this measure.

CO-CHAIR GOLDEN: Okay. So, we go to 2468 and Grace is the presenter. You have a high bar to meet.

MEMBER LEE: I can only try to aspire to decrease my time to Jim's time.

The title measure is adherence to oral diabetes agents for individuals with diabetes. A brief description, at least two prescriptions in a diabetic for a single oral diabetes agent or at least two prescriptions for multiple agents with a diabetes drug class and who have a PDC, as we had previously discussed of 0.8.

So, just to go into some definitions, the numerator is in patients with at least two prescriptions for oral diabetes agents in any diabetes agents in any diabetes

drug class with a PDC of at least 0.8 for at least one diabetes drug class. And the denominator does include 18 years of age or older. These are for the Medicare Part D patients.

As of the beginning of the measurement period with diabetes and at least two prescriptions for a single oral diabetes agent or at least two prescriptions for multiple agents within a diabetes drug class during the measurement period.

So, to move on to the evidence, they --

MEMBER BREEN: I have a question about the measure. I'm confused about the numerator and denominator a little bit. So, does this mean -- sorry, can you just go back up? It said in the numerator patients who receive any type of an oral diabetes drug or any class of diabetes drug. I am just curious where insulin plays into this and if patients who are suddenly switched from orals to

insulin drop out of this and how you would find that out. Because unlike the statins and the hypertensive agents, there is a lot more movement in therapy in the anti-glycemic arena. So, I just didn't understand the details of that. Maybe the developer can comment?

MR. CAMPBELL: Sure, this is Kyle Campbell again from FMQAI.

Due to the nature of the administrative claims data, we weren't able to operationalize a measure for insulin in the same way that we are for the oral medications. So, this measure is limited to adherence to oral diabetes agents.

In this case with this particular measure, a patient only needs to achieve a 0.8 on any of the classes for which they would be taking. So for example, if they had like let's say metformin and they also had glipizide, if they were adherent at 0.8 to their metformin but not adherent to their

glipizide, they would still be counted as a numerator-positive case.

MEMBER BREEN: But my question is someone who is on multiple oral therapies in May. And in June, their provider decides to drop the whole oral shebang and switch to an insulin basal-bolus form. They are going to come out as a numerator fail, if you can exclude patients switching to insulin.

MR. CAMPBELL: Yes, at this time we have not excluded patients with insulin. At this time, we have only limited the inclusion to patients who have two prescriptions for multiple oral agents within the class.

CO-CHAIR GOLDEN: So just to be clear also because this is where you start getting into some brand names and copays, which can be all over the place.

So, if somebody is on metformin and Actos and intermittently misses the Actos because of co-pay issues, as long as they stay

on the metformin, they would meet the requirements?

MR. CAMPBELL: Yes.

CO-CHAIR GOLDEN: Okay.

MEMBER SULLIVAN: I just wanted to clarify. So, this could incent physicians to leave people on oral agents instead of converting them to insulin.

I have a feeling there is going to be a lot more discussion later on in other sections.

CO-CHAIR GOLDEN: All right, so evidence. Let's go back to discussing the evidence, since now we understand the measure.

MEMBER LEE: So the evidence presented included clinical practice guidelines from the ADA from 2013. However, the guidelines did not directly address the topic of medication adherence directly. Then, the developers presented the results, a review of ten studies, which was quite extensive, looking at adherence, including the measure

PDC but other measures, including medication possession ratio and found that all of the studies showed that adherence was associated with improved outcomes, other intermediary hemoglobin A1c or hospitalization rates.

And so based on the algorithm, the workgroup recommended this be a moderate.

CO-CHAIR GOLDEN: I see no comments or questions.

MS. BAL: Voting is open.

(Pause.)

MS. BAL: Okay, the final results are high four; moderate 15. And we will continue on.

MEMBER LEE: So, a performance gap was identified as to the previous two measures. The mean state was 73 percent or 74 percent, with a 15 percent spread, contrasting the physicians a 73 percent mean, with approximately 40 percent spread. So, there was a gap.

CO-CHAIR GOLDEN: Okay, let's

vote.

MS. BAL: Voting is open.

(Pause.)

MS. BAL: Okay, we have high 14; moderate five.

MEMBER LEE: So, priority. The working group felt that this was high priority because based on diabetes morbidity/mortality and based on the studies and their outcomes that were reported.

CO-CHAIR GOLDEN: Time to vote.

MS. BAL: Voting is open.

(Pause.)

MS. BAL: Okay, we have high 13; moderate six.

So reliability. Again, looking at the signal to noise ratio for states, drug plans, and physician groups, reliability testing was met at acceptable ranges for 0.98 for states down to 0.71 for physician groups.

So, it is recommended to be moderate.

CO-CHAIR GOLDEN: I have a question for the developer.

I have heard there are some plans or some activities they actually provide the drugs to patients to try and create adherence. So they either get rid of co-pays or they actually supply the drugs. But I was just curious. Does your mechanism capture that, if that is something that they do or is that so rare not to be even considered?

MR. CAMPBELL: Unfortunately, I don't have any data with regard to that. If it was covered by the plan itself, presumably it would be captured but we would have to investigate that a little further to get any answer to that question and I don't have an answer readily available.

MEMBER HAYDON-GREATTING: Hi, this is Starlin. I do this in our plan and everything, even though it is free to the patient, is still ran through the pharmacy benefit computer process so we can capture it

because we are still collecting the same adherence medication to make sure they are still filling when we only see them every other quarter. So, we want to know that same information.

So, Medicaid and Medicare would still be capturing that.

MEMBER DUDL: The Aetna study speaks to that. Giving the drugs free in a controlled trial only produced about a five, ten percent better improvement, showing that adherence is multifactorial.

CO-CHAIR ROSENZWEIG: Just a clarification. If a patient is on, let's say metformin and is switched to another oral agent and the patient is adherent to the new agent, then they are not -- there is no way that they get dinged for stopping the metformin, is there?

MR. CAMPBELL: That is correct.

CO-CHAIR GOLDEN: All right, I think it is time to open the polls.

MS. BAL: Okay, voting is open.

(Pause.)

MS. BAL: Okay, we have high eight; moderate 11.

MEMBER LEE: So, moving on to validity, they again looked at face validity and had the exact same members saying exactly 78 percent potential threats to validity. Again, we had discussed that they had done extensive research on cash prescriptions at discount pharmacies.

So given that it was face validity, the working group recommended moderate.

CO-CHAIR GOLDEN: Do you want to make a comment or are you -- okay, Tracy.

MEMBER BREEN: I tried to save my comments for the right box. I am learning this whole NQF process.

CO-CHAIR GOLDEN: And when you learn, it will be time to go home.

MEMBER BREEN: I know!

(Laughter.)

MEMBER BREEN: So my concerns about this measure, as opposed to the prior two, relate back to what I said about the change, the dynamic of therapy. And when we think about and again now it becomes very important that this is a Medicare-eligible population because now it is an older population, and the natural history of Type 2 disease is to become, at some point, insulin deficient. And we are going to see large numbers of older patients transitioning off oral therapies onto insulin. And in the interest of simplifying regimen and avoiding poly-pharmacy in older patients, we hope that many of these orals will be discontinued.

So, I am concerned that there is a big chunk of patients who are going to show up as a ding, as a numerator fail, when they are having their regimens appropriately changes or disease changes but the measure hasn't figured out a way to operationalize having patients

who are insulin fallout. So, that is my concern about the specifications of this measure, as opposed to the other two.

MR. CAMPBELL: This is Kyle Campbell from FMQAI. I think that is a really good point and I think we could evaluate an exclusion for patients receiving prescriptions for insulin. We would need to look exactly at how it would be operationalized. But I think operationalizing an exclusion would be feasible; whereas, actually measuring adherence of insulin patients would probably require a different measure and a different algorithm.

MEMBER MILLER: The other thing that may change is not just even the natural history of the disease but also in that specific population, we are also going to see their GFR dropping and medication regimens changing as a result of renal function.

MEMBER HAYDON-GREATTING: Yes, Kyle, I have a question. Some of the dual

eligibles in the Medicaid population in some states are being required to do a four prescription per month limit. So, they are having to take a three-month fill time and spread if they are on 12 drugs. Have you considered this that maybe down the road as one of things that will threaten validity?

PARTICIPANT: Yes, it is happening in the Medicaid population, unfortunately.

MEMBER HAYDON-GREATTING: They can only fill four prescriptions per month. So if they are on 12 chronic meds, they have to -- like January they get four for 90 days, February they get four for 90 days and so forth. And so they are having to --

CO-CHAIR GOLDEN: There are many states, again, this is a Medicare measure. There are many states where Medicaid limits the number of scripts. So, I can tell you that I have had patients in my office I say these are your expensive ones. These are your cheap ones. Pay cash for these. Use your

card for those. So, that gets into the whole cash issue. So, it is a strategy.

MEMBER HAYDON-GREATTING: And it is affecting the dual eligibles in Illinois. I just want to make that point. So, when you are out there looking at the data, which you will be, you might want to consider those states that are making their dual eligibles fall into that four prescription limit.

MR. CAMPBELL: Okay, we will definitely take a look at that. I mean the measure would definitely capture the 90-day supply and give credit for a 90-day supply. But I understand with regard to the four prescription limit, if that is, indeed, the case in some states, for dual eligibles is something that we should be aware of. It is not something we investigated.

CO-CHAIR GOLDEN: Sue.

MEMBER KIRKMAN: But I guess I am confused now because if this measure specifically is for Medicare Part D, so that

wouldn't apply. I mean I understand that if somebody else --

MEMBER HAYDON-GREATTING: There are dual eligibles that some of the services for Medicaid you get your prescription filled through Medicare Part D and so --

MEMBER KIRKMAN: But would they be held to that same four prescription a month rule? It seems like that would be only if Medicaid is their prescription coverage.

MEMBER HAYDON-GREATTING: They are putting them in a managed care. What they are doing in Illinois is they are bidding them out into a managed care population and they are doing some limits because of the funding that is -- because the thought was we could move everybody to Medicare Part D and that would be great for the patient in particular. But because of their living arrangements and getting disability under Medicaid, it is split. So, under their split, they are getting some care.

And if you are a bipolar schizophrenic that is on all these medications and have diabetes because of the medications for all those kind of -- I just wanted to make Kyle aware that some of those might be something that would affect the validity.

Don't do it on Illinois. So, leave our state out of it. It's not good data.

CO-CHAIR GOLDEN: I think another state is don't tread on me but that is neither here nor there.

Anybody else with comments? Bill.

MEMBER CURRY: A procedural question. If we feel strongly that there should be an insulin exclusion, where do we go with the process?

CO-CHAIR GOLDEN: So, we have a measure before us as specified. If you want a change to the specs, you would basically give this a low number or, at some point, reject the measure and ask for revision.

So but we cannot make a revision today. I don't think we can do it on the fly. It is a substantive change. So, if you really feel like you want that to be done, then you would give this a low score and at some point vote no on the -- this is not a must pass -- this is a must pass. Excuse me. It is a must pass.

So, if the group votes against it or doesn't like it, then we can ask for revisions.

MR. MATTKE: One question on that. Soeren Mattke for the developers. In my experience, isn't it the case that even if you go from orals to insulin that initially you keep patients on some dose of oral so that you don't have to dose the insulin side?

CO-CHAIR GOLDEN: In looking at the room, I see lots of faces saying it ain't necessarily so.

MR. MATTKE: Okay.

CO-CHAIR GOLDEN: Are we ready to

vote? Okay.

MS. BAL: Voting is open.

(Pause.)

MS. BAL: Okay, we have high one; moderate four; low nine; insufficient five. And so this does not pass.

CO-CHAIR GOLDEN: So, the committee decided that we would like some revisions or some rethinking.

You got two out of three, guys.

MEMBER KIRKMAN: So, is the issue this insulin issue primarily?

MEMBER BREEN: I think it is a good tool to get at medication adherence. The concern is, I think, there are going to be too many people tagged as being non-adherent because you are not accurately capturing therapeutic change. So, if you just pull out the patients who have some insulin prescription, then you get to the nut of people who are strictly on orals. And then if they are falling out, then you are sure that

it is because of non-adherence, as opposed to the way it is set up now you are not sure what is measuring adherence versus what is trickling in from therapeutic change. Because there is a lot of insulin being used, especially in that population. As the insulin rates go up, if you are not pulling those out, I think it is a dirty number. I don't know how --

MEMBER BREEN: I mean it would be interesting for them to actually look at the data because in the older population, there is a lot of incident diabetes, too. So, it is not necessarily the case that everybody has got long-standing diabetes. And I think if you look at the NHANES data, insulin use is actually lower in older patients than in middle-aged patients.

So it just might be worth them looking at the data and if they find that most people that switch to insulin that do stay on an oral agent or whatever.

CO-CHAIR GOLDEN: So, Helen --

MEMBER KIRKMAN: But that would mean that if they went back and looked at the measure and did that math for us and presented it again and say we have looked at this and we found that it was a negligible difference, I would at least feel more comfortable. But now I don't know what that difference is.

CO-CHAIR GOLDEN: So, Helen, I have a question for you, since we are a new standing committee with a whole different framework. We have only two measures that have gone down but this one went down.

Can the developer contact the committee for just some insight? Do you have any process for that? Is it they are on their own? I am just thinking through here how to make this constructive, if people wanted revisions.

DR. BURSTIN: I think that is a good question. And again, we are just starting on standing committee. This is

somewhat new ground for us. But either way, it is very appropriate at this point if the developer wants to give you the additional analyses that you just raised for you to consider it in a follow-up discussion and you can just stop the analysis here for today and potentially return to it.

Does that sound reasonable, Karen or Lindsey?

CO-CHAIR GOLDEN: Yes, I guess one of my thoughts was after this meeting we all go home. Do you want to have some -- if people are willing to send comments or something, is that something of value or even appropriate?

MS. TIGHE: I will just jump in. Actually, during the commenting period, if a developer wants to investigate this potential exclusion and see whether or not it actually impacts the measure, there will be sufficient time for them to do that. And then we have a call scheduled after the comment period where

they could share that information with you all.

Because the endorsement recommendation, though we ask you to vote here now, we put it out for a comment and then you do have the opportunity to reconsider, based on the comments received and then this potential additional information that you are requesting from the developer.

So this is not necessarily your final, final vote.

CO-CHAIR GOLDEN: So just to get the committee oriented, so as we finish today, everything we did goes on the web. There is a public comment period. The comments come back to us.

And so the CSAC, the higher committee, doesn't see material from this committee until we see the comments and review our previous actions.

PARTICIPANT: Now do we then do an up or down vote after that or do we have to go

through each of these individual --

MS. TIGHE: For this measure, because we stop at validity, we would have to finish the vote on this measure so that we could have that information.

PARTICIPANT: What about the ones we have already approved?

MS. TIGHE: Those, if you see the comments and it causes you to reconsider your initial recommendation, you can request to vote again. But it is not, by any means mandatory.

MEMBER BREEN: So, I guess my question is, is there a way to vote like an asterisk? So it seems like on this one measure we have a very specific question where I assume that the rest of the vote will probably, for the other measures, will be similar to the last two measures. So, I am just -- I don't understand the process.

MS. TIGHE: We will reflect that in the draft report when we post it for

comment that if this one specific issue that the committee is seeking comments on this measure and provide all of that to you. Because this won't be, by any means a final.

CO-CHAIR GOLDEN: So in terms of time line, we will be seeing the results of the comments about June, July?

MEMBER McCOLLISTER-SLIPP: So just a procedural question. So, is this sort of like an FDA advisory committee, where we provide advice and then ultimately NQF and like the master committee or whatever decides whether or not it actually gets approved?

MS. TIGHE: So you do provide these initial recommendations and then revise them, potentially, based on the comments received. We will then put it out for NQF member vote. All of this information is taken as an input to our Consensus Standards Approval Committee. And they have specific criteria that they are evaluating at that point in time. I am looking to Helen now but

it is whether the process was followed, whether there was sufficient consensus achieved at the committee level and at the member vote level. I think potentially a change to evidence in the time between.

Yes, so they are looking for specific things where they may reverse the decision of the standing committee but the work that you are doing now goes forward, unless one of those issues arises.

CO-CHAIR ROSENZWEIG: The measures that have been waived by the developers for a later time, can we fully vote and approve them on a conference call or do we have to wait for another meeting like this?

MS. TIGHE: I will actually let Katie talk about the timeline more but I will say there won't be any more in-person meetings scheduled at this point.

DR. PACE: But your question is the two measures that you didn't get to, you will finish those and that will be included in

the recommendations that go out for comment. So, you will vote on them, yes.

MEMBER KIRKMAN: Sorry. The ones we didn't finish because we didn't get all the way through them here or the two that we haven't talked about at all?

CO-CHAIR GOLDEN: The podiatry ones.

MEMBER KIRKMAN: Okay, so for this measure, I guess I am still confused. We didn't actually get to the end and vote it down. We just stopped the process in the middle. So, is it -- so, we will get more evidence from the developer, presumably, and then complete the vote by phone. Is that right?

MS. JOHNSON: So, this is a must-pass criterion. So right now you have voted down this measure. If the developer chooses to, and I imagine that they will, they can bring you back more information. And when they do that, then you can decide if you want

to revote. If you do revote this criterion and it goes forward, based on what they give you, then we would finish out the voting for this measure and then it would go forward.

CO-CHAIR GOLDEN: Is the voting based on a quorum attendance or is it based upon a percentage of the total committee members?

MS. TIGHE: Yes, so we require a minimum of 75 percent of the committee to vote on a measure. If we get 75 percent of the votes back from the committee on the measure and we realize that the remaining votes could still change the decision, then we will go out and seek the rest of the votes.

MS. STREETER: As far as the time line, over the next several weeks, staff will be preparing a draft report that does summarize all your recommendations from today, as well as from the March 12th call and you evaluate the two APMA measures.

We will be putting the draft

report out for comment the first week of April and that is a 30-day period. We will then be asking you to meet via webinar to review the comments. I believe it is -- I don't remember the exact date. I think it is the third week of May. And then that is when you will review and respond to the comments, decide if you want to reconsider any of the measures. And then in June, that is when we will hold the 15-day NQF member voting.

Eventually, your recommendations will be put forth to CSAC in September.

MEMBER SULLIVAN: So, one of the procedural things. So, I think I understand if CMS does the analysis, they present it to us on the call, the insulin impact is negligible, we could vote on the measure as it stands with that impact being negligible.

Suppose they do the analysis and then insulin impact is not negligible. Is it an option that CMS could say and we have quickly figured out how to do an insulin

exclusion and that they could present it and then we could vote on the revised measure?

MS. JOHNSON: I think so, yes. So, it may not be possible for them to do that but I would think that if they did, you could potentially vote on it.

MS. TIGHE: And I know you were asking if CSAC is the Consensus Standards Approval Committee that I referenced earlier.

CO-CHAIR ROSENZWEIG: What is the earliest time in which measures that we officially approve here can they be implemented?

MS. STREETER: The Board approval period is in October. So, the Board would approve endorsed -- if we are still doing that. Okay. So, it would be October.

CO-CHAIR GOLDEN: Okay. So we finished the measures. Where do you want to go from here? We have a number of items here on the agenda. Time is starting to get tighter.

And so I will leave it to NQF staff to decide which of these items you would like to -- and it might not have been the presenters or the lunch. It could have been the cookies, Tracy.

Which area do you want to go next?

MS. JOHNSON: Okay, so we are not going to do a harmonization discussion this afternoon. One main reason for that is there would be a pretty definite discussion of that after you do the podiatry measures.

CO-CHAIR GOLDEN: If we do the podiatry --

MS. JOHNSON: The podiatry measures, right.

So, we will push off the harmonization discussion until the call on the 12th of March. So, you should already have that on your calendars. We will do all of that then, evaluate the two measures and discuss harmonization at that point.

So, what we were hoping to get

from you guys today is some feedback. So, we have -- this is a pilot project and as part of the pilot and then as part of just regular improvements that we try to do with our process, we have actually instituted a lot of new things that you guys are the first ones to really try out.

So, we were hoping to get from you quantitatively through the survey that we sent you out this morning but also a little more qualitatively this afternoon, any feedback that you want to give us.

And I will just give you the overall pieces that we are interested in feedback about. And I will just let you guys discuss.

So, I want you to think about the orientation call that we provided and the information that we gave you on that, as well as the Steering Committee guidebook that we gave you, and a tool called the "What Good Looks Like." I don't know if you -- so we are

interested did you even take a look at those things. Were they helpful? That sort of thing.

We had two Q and A calls that you had the option to attend. So, we are interested in did you attend and was it helpful to you.

We had workgroup calls. And on the workgroup calls, I think one of the problems that we already know about is we didn't get through all of the measures. So, some of the measures got a very short shrift in the workgroup calls. But let's talk about what you thought about workgroups and what would be most helpful to you, if that is not it.

And then finally, we, as part of your submission materials that we gave you, we did this staff review piece that we kind of placed on top. So, we want your feedback about was that helpful at all or not? If part of it was, which parts? And then as part not

just the staff review part but even the formatting, what it looked like, was it helpful to have it stuck with the submission materials, with the hyperlinks? So it is basically a free for all for you guys to give us input so that we can improve our process.

So with that, I will ask you guys to start.

CO-CHAIR GOLDEN: I just want to make a quick comment as the chair, just that I think that the fact that most of you, or most all of you, have never done this before, I think people did real well. But I think that the workgroup calls, which could have been great, had a certain amount of chaos to them because nobody had done this before. And now that you have been through this meeting, I think they would be a lot more useful.

I hate to say it but the workgroups could have been a really high-risk venture, given the inexperience of everybody in this process.

Sue?

MEMBER KIRKMAN: I just had one comment about I think everything that the staff did, to me, was very helpful. But one thing is the big document with all the information, like by the time of the meeting, it was sort of overwhelming. It is just kind of hard because you have got what the developer put in there and then you still have the staff comments in there. And then you have comments in there. And then you have red highlighted comments in there and they are not always kind of in the same place.

I don't really know what a better way to do it is but maybe if the final document could be sort of pared down to just sort of what we would need at the meeting. I mean I don't think it is good to have it in multiple different documents but that was just my gestalt was the document just sort of becomes very hard to plow through and figure things out.

But I thought the staff comments were helpful in the beginning, I mean especially as a newbie, because it did sort of help orient us.

CO-CHAIR GOLDEN: A question for you, Sue. In terms of the formatting of the document, would it be more useful to have staff and workgroup comments separate from the big document for easier reference or do you want everything all in one big document?

MEMBER McCOLLISTER-SLIPP: I would say yes. I mean, as somebody still has no idea where the workgroup comments were that I was looking for yesterday. I mean I found that really hard to navigate. And if I would have brought my marked up version instead of leaving it at home, it would have been a lot easier.

But I think formatting in one way or another, I don't know if you have a graphic designer or whatever the case may be, could be really helpful, just using color or different

kinds of section headers or something.

MEMBER KIRKMAN: Yes, I don't know whether a different document is better or worse but maybe just paring down something. Like, I don't know that we still needed the staff comments today, for example. So, maybe just cutting them out.

DR. PACE: I guess one of the questions because you were in your small workgroups, but people at this meeting had to review all the measures. So, I guess that was part of our thinking is that those who hadn't reviewed those initially, would that be useful to still have that.

So, I hear what you are saying for the people who had already focused on those. But I guess that was part of the balance. But definitely, we will continue to work with that.

MEMBER SHWIDE-SLAVIN: I liked that you kept the staff comments there. It was helpful today to be able to look at those

as we were going through these.

But I think the placement of the committee comments and all of the information, instead of putting it at the very end where you had to figure out we are talking about reliability now but all the comments are after validity and which goes with what, you had to constantly go back and forth.

So, if they were really placed in the right place, it would mean creating new boxes in the document but I think you need to. I think they need to go there.

MEMBER KEARNS: Yes, I just want to say that once I figured out how the hyperlinks worked, it was really good.

I liked the staff comments because they kind of gave me direction for what I was looking for and what I was doing. So, once I figured out how the hyperlinks worked and I could go back and forth, I found that very helpful because I think the original document just plunked down to me would have been kind

of overwhelming without that.

But yet, as I got more familiar and looked at it more and more, I did go back to the original document more and more to see what I thought about it.

So, I thought that system worked. I would echo about where the comments are placed but that has already been spoken to.

MEMBER BREEN: I agree with the comments about having the staff comments for other measures that I didn't look at so closely. I think that is helpful.

But I think what we are really asking for is two different things; one very robust document that has everything and almost maybe like a cheat sheet, something I can refer to with like just the meat of it. Either just the meat of the staff of comments or just the meat of the workgroup comments that I could almost have side-by-side. So, I am looking at the very detailed measure but then I see kind of right next to it what some

of the key topics are. I think that would move the reading along faster.

And in terms of the learning curve on this, I am very new to this process, I thought the learning curve was incredibly steep and where we were today is very different than where we were yesterday. I almost want to do my workgroup over. I think it would be a much better workgroup. But I don't know how to make that learning curve steeper. Because until you see it in process, I don't know how I could have learned how to do it better, other than to sit through this whole thing yesterday. So that is the challenge for those initial workgroups, for those of us who are new. And I don't have any suggestions how to make that better. But I think that we could have used that time much more effectively if we somehow understood the process.

And I was on the Q and A call kind of listening in and I did all the good things

and I still didn't get it. And I don't think until I sat here I really got it. Not for lack of your trying. Very good job but I had no idea what you were talking about.

(Laughter.)

CO-CHAIR GOLDEN: It may be that when you have -- you know we had a large volume of stuff to do. And it might be, honestly, if you have especially a new group, it might be good to do a measure on a phone call. Have somebody walk through it and actually walk through one, so that everybody can follow what --

MEMBER CURRY: They did do that and it was very helpful.

CO-CHAIR GOLDEN: Okay, maybe I wasn't on that call.

MEMBER CURRY: But it took such a long time, we didn't have time to do the others.

MEMBER BREEN: And we weren't invested in it yet.

CO-CHAIR GOLDEN: Never mind. Ingrid?

MEMBER DUVA: I liked the staff comments and I agree with the executive summary for today. Since we have all done our workgroup, it would be nice to look back and see at a quick glance what the other workgroups had kind of decided.

But from a logistics standpoint, I thought it would be very helpful to have two separate documents where the developer has their document and then the staff, whose comments were helpful and it helped us kind of orient to where we were in the process on the workgroup, especially I was stuck because you can't go into the NQF site twice. I tried to pull up the document twice so that I could be at different points in the document and compare but you can't go in. Your access gets blocked.

So, that is why I thought if we had two -- you know, then you download both

documents or whatever. But if you have the two documents, you can also split screen and put them up side-by-side. And I just thought that would facilitate the work really for me that was my perspective. Because that workgroup call is so important to kind of guide everybody through what you have already done to make it efficient.

And then the staff comments were so helpful, you kind of wanted to have them up so you could say well this is what I saw. This is what the staff saw, just to make it more efficient. And then the executive summary today, I thought that was a great idea.

MEMBER McCOLLISTER-SLIPP: I would agree with that. I found it, and I don't know if it was just me, but I found it really hard to sort of sort through the whole document when you are originally reviewing the measure. And I ended up like wasting, or not wasting but using tons and tons of toner printing

things out so that I could see this is the section, this is this section, this is this section.

So, anyway, I mean I don't know. I mean I hate to kill trees by having you guys print all this stuff and send it out to everybody. But on the other hand, I mean it is much easier to review it sort of logically if you have got one section here, one section here and you can sort of cross-reference back and forth. Because when you go through and answer the questions, you are sort of going from one question to the other to the other. So, you end up flipping sort of.

And for me, the hyperlinks didn't work. So, it was kind of cumbersome, especially if you are doing it while you are traveling and your laptop screen is this big.

MEMBER KEARNS: I really think once the hyperlinks worked, that saved that problem for me because I was having a lot of that. And once I figured out -- and I don't

know what your issue was -- but once I figured that out, it was essentially like having them side-by-side because you can easily go back and forth. And that saved me a lot. A lot.

I didn't figure that out until yesterday.

CO-CHAIR ROSENZWEIG: Yes, I had a big learning curve with respect to that as well. It was not easy.

MS. JOHNSON: So let me ask, what do you mean about they didn't work? Do you mean they actually didn't work or you didn't know how to go back?

MEMBER KEARNS: I didn't know how to go back and forth very well.

MS. JOHNSON: Okay.

MEMBER KEARNS: And I kept ending up in the wrong section in one or the other. And it was confusing until I figured out that they were actually one big document, I was jumping back and forth.

But yesterday I think someone

helped me figure out how to really go toggle back and forth. And then it was like a light bulb went off. It was a lot simpler. Otherwise, it was kind of very I am scrolling way down, I am scrolling way up. Where is it? Where did I start? But they hyperlinks, once I figured it out was like wow, that was a really good idea.

CO-CHAIR ROSENZWEIG: Yes, I kept logging out accidently because of not being able to switch and staying within the document.

MS. JOHNSON: We will send instructions.

CO-CHAIR ROSENZWEIG: I have a couple of questions for the committee as a whole. What about, this may be specifically NQF policy but the process of having the measure developers in the room the entire time we consider the measure, is this cast in -- is this a policy that NQF has fully decided on?

Because to me, that can be a

little intimidating for the group as a whole. And sometimes people may want to -- the committee may want to raise issues related to the overall consideration of the measures in private. Has that been discussed at all? I mean especially at the very end when we decide to vote on the measure. There is a certain amount of inhibition of criticism to a certain extent that is implicit in this process.

MS. TIGHE: Yes, NQF generally is extremely committed to transparency and so all of our documents are available on our website. All of our calls are publicly available. We provide transcripts to every meeting publicly. And so, it is a stance that we have taken as an organization to be as transparent as possible.

That said, I certainly do understand your concerns.

I am kind of looking to Helen now. I am not sure if we really have or even could have a process to get at that, other than

perhaps email amongst the committee.

DR. BURSTIN: Yes, it is really complex. I mean again, so much of our hallmark is transparency. So, we try to be transparent. That being said, there is nothing that would prevent you from, at some point, often at the end of a meeting, going into an executive session and having a chance to freely air any concerns you want to talk about that are not directly related to individual measure evaluation, in which case you want the public and everybody there open. But I do think there may be conversations about specific issues that came up over the meeting where we could do that in executive session.

Is that fair?

MEMBER KIRKMAN: Something related to that, so on the workgroup call, I was actually kind of -- I guess it didn't sink in to me that the developers were on the call until they started to speak up. So, at first,

I was a little taken aback by that.

But once I realized it was, I actually think it was really good because, for example, the APMA measures. I mean we had a lot of criticism of them and I think the developer sent some things in afterwards and there was a little bit of defensiveness on their part as well. But I just think if they hadn't been on the call and if they weren't on the call where we actually approve or don't approve the measure, then I think they are just going to say oh, you know, they just didn't understand it. We got screwed, whatever.

And so I mean I actually think once you sort of realize they are going to be here, I think it is better. I think it is better for us and for them.

And so for example, the osteoporosis measures, I mean one of them didn't go forward. And I mean I don't think it was that hard for --

MEMBER SHWIDE-SLAVIN: Yes, but I also heard them talking that they didn't expect that one to go forward.

That was interesting.

MEMBER KIRKMAN: I mean I did feel like -- are they still on the call? Are people still on the call?

MS. TIGHE: We are not in an executive session.

MEMBER KIRKMAN: Okay. So there was some eye rolling going on for one of the measure developer groups yesterday. But you know I think you just deal with it.

MEMBER BREEN: I think that is okay. I think that is good communication and there is very powerful communication when you have people in the room, which is why I think is more effective than being on that isolated workgroup call.

So, I think for me having developers sitting right in front, being able to speak to the details or even on the phone,

it is very helpful. And I think you learn more. I think it informs their process. I think it is good. I think we can be professionally candid about the things and deficits that we see, that they have to be addressed.

CO-CHAIR GOLDEN: Bill?

MEMBER TAYLOR: I think that there is a difference in how some of the measure developers, though, interacted with the group. And I really appreciated yesterday when you guys, as the chairs, would interrupt their interruptions of the workgroup's conversation.

I think that there are times we need to hear from them but I think there are times that our conversation needs to come first.

So, I think that is important that you set that tone yesterday. And I don't think that we were intimidated by their presence. I think there was one group that was pretty forward in their arguments. And

that is okay but I think by setting the tone that our conversation needs to happen first I think was really important.

CO-CHAIR ROSENZWEIG: Well, I do recall one incidence in which you were presenting a case in which your former boss was sitting over here looking directly at your.

MEMBER BREEN: Yes, good stuff. I like that.

(Laughter.)

CO-CHAIR ROSENZWEIG: And I'm not saying it was --

MEMBER BREEN: No, good stuff. But anyway, this is what we all have to do. And I think the moderators handled it very well. I think groups this big and bulky have to be well moderated.

And I think you said yesterday don't panic if it takes us a gazillion hours on the first measure because we will get through it. So, at the time I was thinking

oh, maybe this needs to be more aggressively moderated but I think we needed to go through that process.

So, thank you for saying it will take so long because it did.

CO-CHAIR ROSENZWEIG: One other issue I would like to bring up is that we are kind of in an age now when there is going to be progressive proliferation of measure sets to a certain extent. And a lot of these are going to be very similar to each other or overlapping in a lot of different ways.

Does the committee have the right to kind of screen the measures? Let's say if there is another measure set that is basically very similar, only with slightly different language to the other set, do we have the ability to kind of screen them and say well, we have already covered this in another measure set?

MS. TIGHE: Yes, so we have moved to standing committees for several reasons.

One is the steep learning curve that you have described in applying the criteria to the measures. Another is the idea that we would like you to take ownership of the portfolio. And so to understand where there are measures that are potentially competing with each other or very highly related to evaluate them independently against the criteria and then really to make an overall assessment of whether we need both of these measures and if not, which one we don't need.

So we do want to make sure, of course, that they both meet the criteria. It wouldn't be the first one in is the one that we keep necessarily. But yes, we would want you to make that determination.

CO-CHAIR ROSENZWEIG: Okay, thank you.

MEMBER DUDL: I'm going to make a statement that I think might make this meeting flow better. We could -- could it be possible that once we are assigned what we are supposed

to do, we get a template from you about the exact items we are to present?

For example, the issue was well, give us a quick review. Well what does a quick review mean? And when we go into each section, how much is that that we would fill that out, send it back to the staff, and like Karen Johnson was fantastically helping me because it was interesting, we use grade for implementation and you don't. And I was going to present grade until she explained how everything works out. And then you send it back.

And I think the whole thing might just go a little quicker. We won't lose our place and that kind of thing.

MEMBER MILLER: I had made up my own little cheat sheet template as I went through this when I started out because for the first measure, I had no idea what I was doing. So, I just thought make up a little cheat sheet template that I could use then for

each of the measures.

I will say, too, just addressing the NQF process, I was very impressed with the amount of data that you provided to us that the measure developers had provided to you. I thought your system was very easy to navigate, once we had that phone call.

I know Reva said at first that it was going to be her default process to go through a measure but I think that was one of the most helpful things.

I also think that going forward for other standing committees, before you have that first call, the orientation call, I think one of the reasons it was easier for me with the orientation call was I had about a little time, like half an hour or so that I looked through the website and the home site of the SharePoint to really look at some of the documents.

So, I think if people have an opportunity to encourage people to take an

opportunity to do that before the orientation call.

MS. TIGHE: I'm going to just jump in on your first comment. If you wouldn't mind sending the document that you used to prepare, that would be really helpful so we can see where it matches and differs from our own.

MEMBER MILLER: Sure. It is very informal. You may be underwhelmed but I will be happy to send it to you.

(Laughter.)

CO-CHAIR ROSENZWEIG: One other thing. Can we assume that prevention of diabetes is within the purview of this committee or is that not within our --

MS. TIGHE: I think it falls under our Population Health Committee but that shouldn't restrict you all from certainly identifying it as an area where we need measures.

DR. BURSTIN: And if you look at

the bubble diagram, that sort of longitudinal framework, it begins with prevention. So if you think that there are important areas to bring in, really like we sort of did with osteoporosis yesterday, I think that is fair game.

We do have a Population Health Committee that does look more at sort of general population screening. But again, if it comes up and it is diabetes, because you are a standing committee, we would likely bring it to you for your input as well.

CO-CHAIR ROSENZWEIG: Another thing is that I was really extremely impressed with the expertise and diversity that was represented in this committee.

I would hope that in the future that as people turn over that they continue to have certainly the diversity that is implied here. I think we always have at least one person who is a methodologist and one person who has diabetes, who can represent people who

have diabetes and also, a pediatric person.

So, I just think that that is very important. I didn't know if when you created, when you selected the members, whether or not you had all of that in mind but I assume that you did.

DR. BURSTIN: Yes, when we do our slates, we very much think about it as a bit of a Noah's Ark. We need some of those, and some of those, and some of those to really get the fully multi-stakeholders. And again, a special thanks to NI. I think it is again, just so important to have the voice of the patient here. The purchaser, the health plans, all of you together makes it a dialogue that, on our own, we just tend not to get quite that rich a dialogue.

MS. TIGHE: Maybe we can make a policy for when you go off the committee, you find your replacement.

(Laughter.)

MEMBER SHWIDE-SLAVIN: I wanted to

make a suggestion that maybe if you had a tutorial available on a 24-hour basis with that sample measure and a lesson how to go through it. That would also be a way that would sort of help the learning curve.

MEMBER McCOLLISTER-SLIPP: Or just maybe like the objective of a particular section. I mean I am glad that everybody else was a little confused as they went through it.

One thing that I think that would be really helpful, and maybe this is more obvious to the others in the room, but what is the intended objective of each measure? Because that will determine a lot about how you think about the evidence presented and the specificity presented.

So, I mean I found that to be confusing because I wasn't sure if I was judging -- if I was creating a measure for a hospital like for uses or a group of physicians who may be, as part of an ACO, if it would be for an individual physician, if it

would be for a health plan. And I understand that once we decide on something and it gets ratified, et cetera, that it can be used for lots of different things. But in terms of your framing of the issue, it would be really helpful to know what the intended objective is.

And I can follow most of the acronyms because I am in D.C. and a nerd and have all the complications but it would be helpful to have at least one reference to what the full acronym means as well.

CO-CHAIR ROSENZWEIG: I think that is an excellent point. Most templates for measure sets usually have like a rationale at the beginning, especially if you require that it be in lay language as well, I think that would be helpful.

MS. JOHNSON: Right. And just to make sure that you get some of our jargon here, we use the term level of analysis to signify if it is a health plan measure versus

a facility measure, versus a clinician measure. And that is actually, if you go back and look at your first page that has that kind of brief description of the measure, that is on there toward the bottom. It is called level of analysis. So, that is where that is located.

MEMBER McCOLLISTER-SLIPP: Maybe if -- and again, maybe this is impossible. Maybe it is an unknown at this point. But I mean I saw that part of it and I certainly thought of that, which was helpful. But like is this a reimbursement issue? Is it an incentive issue? Is it a readmission? I mean, we didn't get into any of that stuff today. But how will this actually be used on the field, once it is ratified? I mean how are we measuring? How is this going to be used and implemented?

MS. TIGHE: I know Helen spoke to this a bit yesterday, so maybe she will speak to it again more. But we are certainly, we do

understand that it is hard to divorce the intended use of the measure from the endorsement of the measure. And so, it is something that we are, as an organization, trying to visit and understand how we can link the use of the measure to the endorsement of the measure. Right now, it is endorsed just generally for all of the purposes under the sun. And then our MAP team takes it and looks at it for use in specific federal programs.

Our Board has certainly indicated that we need to move in this direction. And so that is something that we are continuing to explore.

MEMBER KIRKMAN: Just to make it clear, it is not up to NQF to say this can't be used this way or can be used or should be used this way or whatever. Right? It is just sort of the developer presents it and they sort of say in general this is why they want it and what it might be used for, but after that, it is not really up to you all or we

all.

DR. PACE: Well, we endorse the measure as specified. And under the premise that the measures that we endorse are suitable for accountability applications, in addition to improvement, I mean that is the whole goal of any of these performance measures, but beyond on that -- so, we, from an endorsement standpoint, we would expect it to be used as specified and endorsed. But we don't have control when it gets implemented.

DR. BURSTIN: Although in some ways, it does present a bit of a backstop for end users to be able to push back and say well, this was really endorsed at the health plan level. Why is it being applied at the clinician level? And then to go back to the evidence to say there was an evidence support.

So, I think some of this it can be used in that way. And that is why I think the more we would like to see those decisions about use is really driven by science. And

what you guys have spent a lot of time talking about over the last couple of days of when does a measure logically fit at the clinician level, for example. When is it appropriate for payment, for example, versus quality improvement.

So, more on that to follow. We will definitely be reaching out to you as we kind of think through those next steps.

CO-CHAIR ROSENZWEIG: So, once a measure set is approved and is implemented, the measure developer doesn't have to come back to you with outcomes until they present the measure again for renewal?

DR. BURSTIN: That is correct. They come back in three-year maintenance.

Now, we do have something called our ad hoc process, which at any point during the time when a measure is endorsed, if there is either a change in the evidence that would substantially affect the measure or any evidence of unintended consequences, we would

immediately re-review the measure. So, we do have that as our sort of backup, particularly for changes in evidence, which happen, as we all know. That is why we have delayed the lipid measures a fair amount.

So, we recognize that it is just not static to assume that this won't change. And three years was the number we picked, frankly, because it mirrors the usual periodicity of guideline development as well. So, it seemed like the logical number. But if something happens more acutely or there is evidence of unintended consequences, we will review it again.

CO-CHAIR ROSENZWEIG: Any other comments?

MEMBER KEARNS: I just had another comment on a different topic, we are okay with that.

I found that I had a hard time finding the SharePoint site. What I wanted was to be able to log in to NQF, go to my

dashboard, boom there is the SharePoint. I click on it. Instead of that, I had to kind of keep that link separate because I didn't really know how to find it. I looked all over.

So that is kind of a technical thing. I wanted to have my dashboard really drive everything and I don't know what the answer to that is.

The second comment is I thought the travel arrangements were fantastic. It was very easy to make the reservations for the flight, the hotel. All of that was very easy from my end. Thank you.

MEMBER BREEN: I enjoyed the cookies. Thank you very much.

(Laughter.)

CO-CHAIR ROSENZWEIG: Yes, Bill?

MEMBER TAYLOR: I also really appreciate how much staff support there was that I felt very well taken care of. I also felt very lost at the beginning and sort of

mired in jargon that I was unfamiliar with about what the process was. I'm not sure how to do it better to get us up to speed because I think we all got there but the start is hard.

But the amount of help, whenever I needed it, was unbelievable. And the responsiveness of anybody that I got in touch with I'm looking especially at the Board was just phenomenal. So, I deeply appreciate that.

I also appreciated that we started with the attention to conflict of interest about all of us and were explicit about that.

I am still a little lost about I have a sense of all of us, of where we are and where we come from. I have less of a sense of NQF and where does it come from. And now that we have all been in the midst of it, maybe everybody else is more sophisticated about that. But where does the money flow? Who paid for our salmon and all of that and our

plane tickets and so on is now of more interest to me, having participated and being a part of this.

But I am extremely impressed by how well-run the whole operation is from what I have seen of it.

DR. BURSTIN: Well, thank you. We have great staff, certainly, and appreciate the comments.

At this point, the overwhelming majority of all of our endorsement work is funded by the federal government. It is funded through a contract with CMS. We do still have some dollars from foundations and from our membership dues, from our organizational members but, overwhelmingly, it is funded through CMS.

So, I think we actually buy the cookies and things. The salmon is ours, which is why it is all family style. When we moved to this building, we intentionally set up this room so that we could avoid all the hotel

lunches for $45 for a sandwich and so we can use local caterers and get good meals and good deals. And it is just a lot more comfortable.

So, thanks.

MEMBER McCOLLISTER-SLIPP: I guess Bill has left, so we are not going to get the brandy.

(Laughter.)

MEMBER McCOLLISTER-SLIPP: But just from a philosophical -- I mean this is more of a philosophical -- maybe this isn't the best form for it but from a patient perspective who thinks that we have a long way to go in terms of establishing good meaningful quality measures, especially for those of us with Type 1, I found it slightly depressing when I looked at all the documentation required. And I know that sounds kind of silly because you do want this to be science-based and you do want -- I mean I understand the need for all of it. But I think and maybe this will be answered in the incubator thing

that you were referencing yesterday, Helen, I think it would be great and I feel some sort of a need of this committee or some forum to be able to articulate a path forward for new measures. Because I mean hemoglobin A1c is based on the science from what, 40 years ago. It was endorsed 20 years ago. And we need to have a way of encouraging, directing, incentivizing, whatever, the development of new measures that take into account all of the things that we have learned with the great science that has been done, with all the studies that we do have, and with new technologies like continuous glucose monitors as a mechanism for doing different types of measures.

I don't know what the answer is.

DR. BURSTIN: It's a great point. And actually one other issue I will raise because this is a committee where we do have periodic opportunities for submission. The other thing is we often don't get, for

example, some of the measures some of you and your health systems are using on the ground that you have found incredibly valuable or that patient groups have come up with. So, the question is really how do we work with those.

And so that is part of the idea of the incubator as well. So, for example, if Kaiser or UNC has a great measure that works perfectly well in their system, can we bring that to the incubator, bring the right resources together to create, to pull it into a national standard and get all the rest of that work done. Because we do recognize there is a lot of innovation out there that is just not coming our way and some of that is because of the burden.

And I see Patricia's card is going up. There is a fair amount of burden associated with being a measure developer and submitting to NQF, which is why we want to try to create those partnerships.

DR. PACE: But I think that leads to a question. Would you drop any of the criteria that we have, thinking about these being used in an accountability application? Because it is a question that we get about, you know our discussion about the burden.

And so basically, to reduce -- sometimes developers provide more information than we ask for. If you looked at our examples of what good looks like, we really do have examples of concisely providing the information we want. So, that is part of it.

But in general, in order to reduce what we ask for submission is to maybe eliminate some of our criteria.

So, I am just curious after having gone through this if there is something that you think is well, maybe not that necessary that we should at least be thinking about.

CO-CHAIR ROSENZWEIG: Yes, Bob?

MEMBER BAILEY: I would advocate that the criteria that you have are good.

Coming from a consumer standpoint as well, very often implementing these measures at a large scale are not unlike a controlled clinical trial, where it needs to go through an IRB to make sure that the evidence suggests that there may be a benefit. You talk about the potential threats, the unintended harms. It is very comforting to know that this is the case. But I can't think of anything that I would eliminate here.

CO-CHAIR ROSENZWEIG: Janice?

MEMBER MILLER: I would like to just say that I really liked that this morning when we were talking about the diabetes education for the foot exam, we didn't have a really strong body of evidence. But I really liked that we could say it was insufficient evidence with an exclusion. And I just think that that is also the path to go for overall diabetes education measures, that it may be a long time until we have a very homogeneous type of education that we can evaluate.

But if we can have something like that, it is very intuitive but we just don't have evidence. So, I think if we could do that with an exclusion, that would be superb.

MEMBER McCOLLISTER-SLIPP: Yes, or some way of just appending things, so that it is not just an up or a down. I mean, some of this stuff is like yes, I want everyone to have a retinal exam but I mean that is how I probably kept a large part of the sight of my right eye, just a regular retinal exam.

But we need to come up with better ways of tracking it or what exactly -- just some of those questions, I think if there is a way of incorporating either nuance or further recommendation or perhaps even a provisional recommendation or a provisional approval with a recommendation for additional studies or additional clarification because I don't want to be a progress to encouraging and incentivizing good care. I don't want to get in the way of that. But at the same time, I

would be a little uncomfortable on some of the stuff just pushing forward with the way that is stipulated.

CO-CHAIR ROSENZWEIG: Patricia, did you have a comment?

MEMBER McDERMOTT: I was just going to comment on you require a great deal of due diligence in order to present a measure. Within Aetna, we have developed lots of measures that are truly Aetna-defined. We are using clinical evidence and we are using valid reasons and good code.

One of the other things that didn't come up here was maintenance of code. Now, the measures that we have talked about don't have a whole lot of -- if you are looking at the class of a drug, you don't have to worry about individual NDC codes or CPT-4 codes for example. But that is another thing that is important is the maintenance of those measures.

But going back to the concept of

submitting a measure, I mean I need to have a whole PhD staff in order to present all the documentation that would demonstrate to all of you that the things that we have done are valid.

We have done our due diligence to look at variability. We have put them out in the market. We know how our doctors respond. It has been positive. We know we are showing differences. But for me to bring those things and present them here, again, I need a staff to be able to do that.

So, I don't know how you change that paradigm. The only the other thing that I would say about some of the measures we have talked about today that are only for a specific population but they are obviously generalizable to a very large population. The only thing that I would encourage is that when we do get a measure that we know has been only researched on a small population but is valid for that larger population, that we try to

figure out a way to get that in the specification. Because, unfortunately, we also have people that if we take the measure and we take it slightly -- it is only supposed to be for this or it is only supposed to be for that, if we add that pregnancy exclusion, for example, or say if it is 40 and above, then the measure is valid. But then we have deviated from the NQF standard that has been approved, so we get a slap on the hand.

So, it is an interesting paradigm what is going on within measurement. And you are a very powerful organization, very powerful.

MEMBER LEE: I had a question on how the developers actually developed the measurement. Is there sort of a letter of intent or is there an intermediary feedback process, for example, the measure I had, in terms of getting voted on, the insulin was sort of the point at which it got halted and perhaps could have pre-screening or some sort

of mechanism could speed along the efficiency of which we are able to evaluate these measures?

MS. TIGHE: Yes, so I can speak briefly to that. And we are trying to figure out ways where we can get the right measures that are going to meet our criteria into our process. And so part of this we held an event in September. It was collaborative. It involved our developer colleagues, federal partners and our staff. And we asked that exact question. How can we get this feedback while the measure is being developed, so that when it comes to NQF for endorsement, it is meeting our criteria and then can be put out into use pretty quickly.

So, Karen actually probably can speak to us because she is working most directly with the group that is still focused on this effort. But one of the things that they are trying to address is by bringing the patient involvement into the development

process in a more significant way, so that they are getting this kind of feedback as they are developing the measure.

Some other things that we are working on are providing a way for them to kind of check in with NQF about whether or not their testing plans are appropriate. That actually really was less of an issue for the measures today. But we do often get measures that meet the importance criteria and that are considered reliable by our committees.

So, we are trying to work with them upstream as they are developing the measures, rather than having our process be at the very end once all the development dollars are spent.

MEMBER BREEN: I mean I would hate to add anything more to this agenda because it is a very ambitious agenda that we have just got there. But it might be interesting to get a little preview of things that are in the works like that while you have this critical

mass of expertise in the room. And again, you don't want to sway things too much but it might be useful for people to be able to pitch an idea or whatever that they are working on.

DR. PACE: And each of the developers has their process and advisory committees. But oftentimes, things are raised when it comes to NQF that, for one reason or another, weren't raised in their development process. So, it is always interesting and we have toyed with different ideas of how to have earlier input but it has been challenging.

CO-CHAIR ROSENZWEIG: I mean as a Steering Committee, I think it would be nice if we had like time to reserve or be able to reserve a certain amount of time to discuss potential candidate measures that we would like to see come our way in the future.

Now, you have this wonderful newsletter that you sent out that I get all the time. And at least in theory you can include candidate, potential candidate

measures that NQF itself is interested in considering.

Sue.

MEMBER KIRKMAN: I've almost forgotten what I wanted to say. But I agree with Anna that we do want new measures but I am not sure that we just want more measures.

So again, I think we also need to be thinking about how to sort of drop measures. And I actually think for accountability measures, if things are going to get publicly reported, I think it does need to have a high bar and be very rigorous. So I actually don't think I would change too much about -- I certainly wouldn't drop any of the criteria. Because I think when you are talking about something that is going to be out there for everybody to see, you don't want kind of a bad measure.

And similarly, we don't want to come to this meeting and review 12 measures and only approve two of them. So, I think you

need to keep the quality pretty high.

But the other thing I wanted to say was just I would, and maybe you all have already done this, but I would just love to continue this discussion of how do you move increasingly toward individualization of care and yet be able to measure quality, which kind of by definition you have to sort of bucket large groups of people into the same bucket and say that the same thing should happen to them. And I just think that that tension is just going to get greater and greater with time.

MEMBER McCOLLISTER-SLIPP: And as a patient, I mean I think that is really, really critical. I mean, and I don't know what the answer is. You guys are smarter than me when it comes to this stuff but we have got to come up with a way to incorporate into policy and incentives outliers and the need for individuality and care. Otherwise, it creates all sorts of unintended consequences

that are pretty negative.

DR. PACE: So, one way that people talk about that has maybe more merit is really focusing on outcomes and including patient reported outcomes. Because then you are not so focused on process. You are focused on the provider using their best process to get the best results.

Now of course, we have risk adjustment issues but that is one advantage of outcomes because it frees you up from having to precisely specify the process and expect that everyone has to deliver the same process the same way to every patient.

And again, I think we should have more of those conversations in this group. But I think that is certainly one advantage of outcomes and patient-reported outcomes that we have a lot of need for.

CO-CHAIR ROSENZWEIG: Yes, Jessie?

MEMBER SULLIVAN: I just wanted to follow-up on that. I do think that that

underscores, the patient-centeredness underscores the contradiction we were talking about when we were looking at the statin measure. And there is a certain bluntness to our measure. And I think there is no question but that it is directionally correct to look for adherence in a Medicare population who have been started on statins. That is the bigger problem.

But when you get down to the individual physician level, and okay, it is not specified for the individual physician but that doesn't mean it won't be applied that way. And when you get down to that level, the fact that someone has two patients who should be excluded, they have done the right thing, it is just extremely irksome and it makes it very hard to --

So, I think it is a contradiction. And one of the things that I think might be helpful is if in publishing the measures as they are, some of that could be teased out in

a comment. And it isn't. I mean, it is buried in the specs but we know the specs aren't always going to be followed exactly.

So, if in the specs it says this was built for Medicare and the issues that might apply if you use this in a non-Medicare population are, statins aren't always indicated in younger people. They are contraindicated in pregnancy. Some people with low incomes may be buying their prescriptions.

So if just some of the things we identified were in a signing statement, it would alert people if they start using the measures in different ways.

MEMBER McCOLLISTER-SLIPP: Or perhaps even statements about inherent biases within population groups. So, I mean again, not to harm on erythropoietins but I want to take a erythropoietin and have a hemoglobin over 11 because I want to go to the gym five times a week, not like struggle to get there

once. And the population that was measured for that particular outcome was probably not looking at getting to the gym five times a week.

So, there needs to be some degree of consideration for what is the population that was studied to support this measure versus what are the real life applications that are going to happen based on the fact that we have endorsed this measure?

You know, similar with hemoglobin A1c in some respects.

MEMBER BREEN: I think what we are all trying to say is rather than just that up-down vote, we would like the outside people to know the very robust conversations what we have had with some summary statement.

I have always found that all these expert groups put out opinion statements. But what I find interesting is not the statement but the back story, the discussion that went on there when people either write an addending

article or something to say this is the discussion. That is where the real interesting stuff comes out. So, I think that having a comment field or just some group comment, yes we have passed this measure and this is some of the concern that the committee had would be an interesting -- but I don't know if that happens right now at all.

So, I think that is what we are looking for.

MS. JOHNSON: That does happen to some extent. We usually have pretty detailed notes of what happened in the committee because that is our way of kind of showing posterity what was discussed.

DR. PACE: But I think this is consistent with other things that we have been talking about and thinking about implementation guidance, something that has come up very similar to these suggestions in our risk adjustment panel is really being explicit about how the measure was -- the

endorsed specifications, what patient population, what setting, what level of analysis, risks of misuse.

So, I think you will see some of those recommendations that maybe resonate, even though that is specific to risk adjustment. Some of their recommendations really have applicability to some of the things you are also mentioning.

MEMBER DUDL: I want to support your idea of going with outcomes. If we went for heart attacks and strokes in the diabetic population, some groups will use titration. Some will use initiation. Some will use adherence to ACEIs or statins or whatever works.

Far more powerful, far more driving, and far more simplistic, yes, it can't be used at the individual level, this is the health plan level, but it could be the single most powerful driver and it eliminates all this issue of process. I think you are

right on and I would keep feeding that back to the people who give you or ask you advice or give you measures.

CO-CHAIR ROSENZWEIG: Okay. Any other comments?

MEMBER DUVA: I just wanted to give quick feedback, jumping back to the evidence. We had a long conversation about evidence. I thought the algorithms that you gave us were very helpful to keep us all on the same page, so I just wanted to say that.

The other thing I thought was interesting and I am not a psychologist or group process expert but I really think that because we started with the measures that have been in use, that our assessment of the evidence was less critical for the measures that have already been in use that we are kind of used to now, hemoglobin A1c that has been reported.

And then when we were identifying the evidence and its exact applicability to

the new measures, we were a little bit harder on those measures. And I am not saying that is a good bad thing. I just say it mostly because Sue keeps bringing up like when are measures going to get dropped. When have we kind of been there done that with a measure? I think it is going -- if that is the group thing, just that is a natural thing way to think about it might be kind of hard. I don't know.

DR. PACE: Along those lines, I am just going to throw out a question to all of you because we had a little bit of a surprise.

So, along these lines of retiring some measures, moving on, we had that situation yesterday with continuing to report on whether hemoglobin A1c test was ordered once a year or not just ordered but given once a year, when we had that embedded in measures that were actually about the results.

So, I guess we would just like to hear more of your thinking about why you felt

that was necessary to continue that measure, especially in light of some of the conversation we have just had.

MEMBER DUVA: Just real quick I want to say that my impression was that we passed it here but it would come back up when we talked about harmonization and which measures were redundant. Is that not coming back up again?

Because I feel like we didn't necessarily finish the conversation.

MS. JOHNSON: I kind of wondered if maybe --

MEMBER DUVA: That was my impression.

MS. JOHNSON: Yes, so the only thing that you would be looking at in a harmonization throwing stuff out later on would be a head to head competing measure.

So, am I not saying that right, Karen?

DR. PACE: That is exactly right.

I was just going to say we hadn't really thought of harmonization or competing measures in that way. So, it is just something we need to kind of maybe rethink about. But that is a good observation.

We have thought of it more if there was another measure specifically about measuring hemoglobin A1c, not the issue of that is actually embedded in another measure. So, that is an excellent observation that we need to think of in our process.

MEMBER BREEN: But I think those measures were also different. I mean the measuring of an A1c is a pure process measure. Do you do just the basic stuff? The other two things told you different things about the care of your patient. So, I don't think that those are necessarily redundant. I mean I don't need to rehash yesterday's discussion but I think in my mind, those were actually similar but different issues.

MEMBER DUVA: True, but we didn't

really finish the conversation of putting them altogether. You know, we did a little bit more today where we said are we looking at these together or separate? Because there are some things that you might consider. And maybe it is that they are different enough that you keep the process measure when you have got an outcome measure but I didn't feel like we finished that conversation.

MEMBER BREEN: We did not finish the conversation.

DR. PACE: No, that is helpful to us. That is why we wanted to ask because we obviously haven't framed it that way and we need to think about it.

MEMBER BAILEY: I think some of the other considerations also were the sample size, the availability of the data. And I think Jessie had provided the example yesterday in terms of so you have a sample of 400 patients, for instance, that may or may not involve a specific provider. So, they are

not able to give feedback first of all how they are performing but more importantly, develop actionable lists for people to act upon. Where if you had a larger sample, like administrative claims, then you can go back to Dr. Bailey and say, okay, you have got these 15 people who haven't values within the last year.

MEMBER SULLIVAN: If I could just underscore that about how we use these measures in our Medicaid health plan.

So, when we have -- we used to have more leeway than we have now to create our own measures as Aetna does. But now the Medicaid Office of the Inspector General of New York would get on us if we used non-NQF-endorsed measures or measures that haven't been endorsed. So, our hands are a little bit more tied in inventing our own measures.

And the screening measure, we use it as a composite. We look at nephro, eye exam, and A1c and LDL. We look at them

together and we create a denominator of every patient in the health plan who meets the criteria for diabetes, regardless of time of enrollment. And then we send those names to every primary care doctor in the health plan. So, we are putting in front of them a list of all their patients with diabetes.

If we were just restricted to the chart review measure, we wouldn't have a measure that allowed us to do the outreach around the broader population that has diabetes. So, it would really harm our quality improvement work if that measure -- I mean the way we use it, it could be composite of those four measures wouldn't hurt us. But that denominator of the patients who just have diabetes and have not been in for service is really critical.

So, you would do harm to us, the work we do if you took that measure away.

CO-CHAIR ROSENZWEIG: Sue?

MEMBER KIRKMAN: So, I understand

that about this measure, although I was one of the people, maybe I was the only person that voted to drop it. But I do think we -- and I think it is really hard to drop things but I think we are going to have to continue to wrestle with this because there are so many performance measures in diabetes and there are so many -- and the primary care people in the room can say this. I mean there are so many performance measures about everything and we don't want to just keep adding more and more and more.

And I brought up the thing yesterday about you know so if you are collecting this data, then you are not collecting something else. And I agree it is an easy one to collect but I just think we do have to be more open to dropping measures, even though it is hard to do. Because we all kind of think about well, there is some benefit and that is going to be true of pretty much any of them.

MEMBER McCOLLISTER-SLIPP: That kind of gets to my point, though. And I completely understand what you are saying and the last thing we want to do is add additional burdens for PCPs and endos because that is a huge issue. But I mean if the intent is to do public reporting for consumers to eventually be able to judge, that is going to be one of the things that they specifically look for, that they will know to look for; whereas, statin, ACEI versus ARB, or whatever, might be a little bit more abstract.

I mean we now have gotten to a point, which is very different than we were five, seven years ago, where people kind of get that as a measure that they need to go for.

So if they are looking at a quick cheat sheet of is this doctor any good and he has a low score for testing for hemoglobin A1c, then that is a pretty good thing that somebody who is not nerdy enough --

MEMBER KIRKMAN: Yes, but it is only going to screen out about seven percent of the doctors because the adherence is something like 93 percent or something.

And you could probably screen people on other measures. I mean it is probably concordant with other things.

DR. BURSTIN: I think part of what we are also -- I think what we also see is that committees tend to be hesitant to take something away when there is not something else better there. And that has come up a fair amount.

So for example, I chair the Quality Measures Workgroup for the Health IT Policy Committee. And I was going back and forth yesterday with my friends at ONC saying so, what is up? Anything new on the EHR space for diabetes? And they have looked at, for example, doing delta measures of where you start an A1c and where you wound up. There are so many issues of figuring out where the

baseline is, which follow-up, et cetera.

So, they are exploring all these new ideas, which is why I think it would be wonderful to have this group give input. One thing to start thinking about is could you do something like time and therapeutic range in a given year. So, you actually get a more meaningful number which you can do with electronic data that you can't do with this sort of forced cut point.

So, I think the more time, now that you are a standing committee, we can actually work with you on sort of the future, I think will make it easier to let go of the past. Because I think it is kind of hard to let go of what you have when there is nothing else there, particularly to Anna's point, where people are wanting some meaningful information to use.

You know the health insurance exchanges have been trying to put data into measures for people with chronic illness.

There is just not a lot available, as people are searching.

MEMBER TAYLOR: The idea of getting out not only the final recommendation that comes out of the standing committee but something about the rationale for how we got there would actually help a lot for issues like getting rid of measures that we think are outdated.

I think there was a discussion yesterday sort of offline about the foot care measure, where we talked about gee, if you vote that down, people will misinterpret it and think that foot care is not important, as opposed to it is already being accomplished at a high level and so on.

But if there were some place prominently to say this is why we did what we did so that anybody -- I'm sure most of the people who use this are not going to be nerdy enough to want to go find those details. But if a discussion is happening somewhere about

how are we going to apply this or what did they mean when they did that, if it was easily accessible, it would make us feel less worried about voting certain ways because there were unintended messages that we didn't intend to deliver and so on. We could say what the message is.

You know we talked about excluding pregnant women with statins but we didn't worry so much because it was Medicare or whatever it is. If we could say that somewhere that people could get access to, it might help us to try to actually adhere to the evidence and say something that we could feel proud of.

MS. JOHNSON: So, this may be something that we will ask you to help us with. We actually do include those kinds of details in our current printed written reports but they are, you have to get into the weeds to see that kind of stuff. And we are trying. And another thing that we are experimenting

with is changing our report format somewhat to include all that detail but also include more high-level things.

So usually, we provide those draft reports to the committees and ask them to look at them. And a lot of times, I mean you guys are busy, but maybe the first one that we try you might -- we might ask you to spend a little bit more time than you might otherwise do, since it will be our first one and see if we are getting to where you think that would be. And we haven't figured out what that is going to look like yet.

DR. BURSTIN: Actually, building on Karen's comment, it might be interesting to actually send you one measure, the very intensive display we usually put in and the technical review reports. And then this idea we have been having of how you try to get exactly at what you just said. Almost an executive summary of the measure, to Anna's point earlier, why it is important. What does

this impact? What are the issues discussed? Karen's laughing because she doesn't believe you can do this in a couple of paragraphs. But I think it would be really good for us to be able to bounce that kind of thing off you, even before we do all whatever it is, 20 measures and be glad you are not on cardiovascular safety because they have about 60 measures. So, this isn't so bad.

CO-CHAIR ROSENZWEIG: Anyone else? I think people still have their things up without really wanting to comment any further.

Okay, well than you very much. I think this was an extremely productive meeting.

MS. TIGHE: Operator, if there anyone on the line who would like to provide any public comment.

OPERATOR: Okay, if you would like to ask a question or make a comment, please press * then the number 1 on your telephone keypad.

And there are no comments at this time.

DR. BURSTIN: Thank you to everybody. We realize this is a huge investment of your time. And especially thanks to Jamie and to Bill for guiding us so effectively.

So, thank you and safe travels home.

CO-CHAIR ROSENZWEIG: Thank you.

MS. STREETER: And also just quickly as for next steps, we will be sending you the call information for the March 12th meeting, where we will review 0416 and 0417, the two APMA measures.

And I think that will be it. Oh, and if you could, please complete the surveys in either SurveyMonkey or we have paper copies as well. That would be really helpful.

Thank you all.

(Whereupon, the above-entitled matter went off the record at 2:40 p.m.)

N

O

T

U

C E R T I F I C A T E

This is to certify that the foregoing transcript

In the matter of: Endocrine Measure Endorsement

Before: NQF

Date: 02-27-14

Place: Washington, DC

was duly recorded and accurately transcribed under my direction; further, that said transcript is a true and accurate record of the proceedings.

Neal R Gross

Court Reporter