NATIONAL QUALITY FORUM

+ + + + +

PERINATAL AND REPRODUCTIVE HEALTH STANDING COMMITTEE

+ + + + +

TUESDAY
MAY 3, 2016

+ + + + +

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., Kimberly Gregory and Carol Sakala, Co-Chairs, presiding.

PRESENT:

KIMBERLY GREGORY, MD, MPH, Vice Chair, Women's Healthcare Quality & Performance Improvement; Department OB/GYN, Cedars Sinai Medical Center, Co-Chair
CAROL SAKALA, PhD, MSPH, Director of Childbirth Connection Programs, National Partnership for Women & Families, Co-Chair
J. MATTHEW AUSTIN, PhD, Faculty, Johns Hopkins School of Medicine
JENNIFER BAILIT, MD, MPH, Clinical Director, Family Service Line, MetroHealth Medical Center
AMY BELL, MSN, RNC-OB, NEA-BC, CPHQ, Outcomes Specialist, Carolinas HealthCare System
TRACY FLANAGAN, MD, Director of Women's Health and Chair of the Obstetrics and Gynecology Chiefs, Kaiser Permanente
GREGORY GOYERT, MD, Division Head, Maternal-Fetal Medicine, Women's Health Services,

Henry Ford Health System

ASHLEY HIRAI, PhD, Senior Scientist, Maternal and Child Health Bureau, Health Resources and Services Administration
MAMBARAMBATH JALEEL, MD, Associate Professor of Pediatrics; Medical Director, Parkland NICU, University of Texas, Southwestern Medical Center
DIANA R. JOLLES, CNM, MS, PhD, Quality Chair, American College of Nurse-Midwives
JOHN KEATS, MD, CPE, CPPS, FACOG, FAAPL, Senior Medical Director, Cigna
DEBORAH KILDAY, MSN, RN, Senior Performance Partner, Premier Inc.
NANCY LOWE, CNM, PhD, FACNM, FAAN, Professor, University of Colorado-Denver College of Nursing
SARAH McNEIL, MD, Core Faculty and Director, Contra Costa Medical Center
JENNIFER MOORE, PhD, RN, Executive Director, Institute for Medicaid Innovation
KRISTI NELSON, MBA, BSN, Women and Newborns Clinical Program Manager, Intermountain Healthcare
JULIET M. NEVINS, MD, MPA, Medical Director, Aetna
SHEILA OWENS-COLLINS, MD, MPH, MBA, Chief Medical Officer, Johns Hopkins Healthcare, LLC
CYNTHIA PELLEGRINI, Senior Vice President, Public Policy & Government Affairs, March of Dimes
DIANA E. RAMOS, MD, MPH, FACOG, Medical Director, Reproductive Health, Los Angeles County Public Health Department
NAOMI SCHAPIRO, RN, PhD, CPNP, Professor of Clinical Family Health Care Nursing, Step 2, School of Nursing, University of California-San Francisco
MARISA "MIMI" SPALDING, JD, MPH, Policy Analyst, National Health Law Program
KAREN SHEA, RN, MSN, Vice President, Maternal Child Services, Anthem, Inc.

SINDHU SRINIVAS, MD, MSCE, Associate Professor and Vice Chair, Quality, Obstetrics and Gynecology, University of Pennsylvania Health System and Perelman School of Medicine

RAJAN WADHAWAN, MD, MMM, CPE, FAAP, Chief Medical Officer and Medical Director of Neonatology, Florida Hospital for Children

CAROLYN WESTHOFF, MD, Msc, Director of Family Planning and Preventive Services, Sarah Billinghurst Solomon Professor of Reproductive Health, Columbia University

JANET YOUNG, MD, FACEP, Carilion Clinic, Virginia Tech-Carilion School of Medicine

NQF STAFF:

ELISA MUNTHALI, MPH, Vice President, Quality Measurement

NADINE ALLEN, Project Manager

KAITLYNN ROBINSON-ECTOR, Project Analyst

SUZANNE THEBERGE, MPH, Senior Project Manager

REVA WINKLER, MD, MPH, Senior Director

ALSO PRESENT:

ERIKA EDWARDS, PhD, MPH, Vermont Oxford Network

CORINNA HABERLAND, MD, Stanford University *

SCOTT A. LORCH, MD, MSCE, The Children's Hospital of Philadelphia *

ELLIOTT MAIN, MD, The Joint Commission

CELESTE MILTON, MPH, BSN, RN, The Joint Commission

PAMELA OWENS, PhD, Agency for Healthcare Research and Quality *

GUSTAVO SAN ROMAN, MD, FACOG, Birthrisk.com, LLC

ANNE SANTA-DONATO, MSN, RNC, Association of Women's Health, Obstetric and Neonatal Nurses

STEPHEN P. SCHMALTZ, MPH, PhD, The Joint Commission *

ANN WATT, MBA, The Joint Commission

*** present by teleconference**

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

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

8:08 a.m.

CO-CHAIR SAKALA: Good morning, everyone. Good to see you again.

I would like to give a little recap of yesterday. We were very hard-working -- I heard people say they were very tired at the end of the day -- thoughtful, and efficient, I think, in what we did.

We covered 15 measures. Ten of them were maintenance measures and five are new. Of the new measures, we recommended three for endorsement, including a significant addition of robust, new contraceptive measures to the corpus of NQF-endorsed measures, pending approval down the line.

One was considered consensus not reached. That is the thermal condition in low-birthweight babies, and one of the new measures was not recommended for endorsement, structural attributes of facilities with high-risk women.

We did 10 maintenance measures. One

was not recommended, frequency of ongoing prenatal care. One consensus was not reached, prenatal and postpartum care. Eight passed, and five 100-percent support, the remaining with 88 to 96 percent. So, I think that was a good day's work. And today we have eight more measures, two that are new and six that are maintenance.

Okay. So, let us begin. Today we will include The Joint Commission core set. We have developers here, Celeste and Elliott.

The first measure will be a third measure related to the last two we considered yesterday. That is 1731, PC-04, Health Care-Associated Bloodstream Infections in Newborns.

So, we will ask you to give a little intro to that. Thank you.

And then -- excuse me -- no one is recused, and Janet, Greg, and Florencia, but is she here today? Okay. And Jennifer Moore is recused on this. So, this will be Janet and Greg as discussants.

MS. MILTON: Okay. Good morning.

Yes, the measure that we are going to discuss is PC-04, Health Care-Associated Bloodstream Infections in Newborns. The denominator is comprised of live newborns with two included populations, newborns that are 500 to 1499 grams, newborns that are greater than 1500 grams that had one of the following: they would have had to have experienced death, major surgical procedure, on a mechanical ventilator, or received them as they transferred to the hospital. These are the babies that we would consider more at high risk. And then, if they were in the numerator population, it would mean that they had a newborn bacteremia or septicemia. And that is what would be evaluated in this measure.

It is an outcome measure. So, we do have a risk-adjustment model. The goal is to have a lower rate as noted improvement for this measure.

CO-CHAIR SAKALA: Okay. Thank you.

So, can we have our discussants lead off, starting with the evidence?

MEMBER GOYERT: Sure. As far as the evidence, the developers provided no new evidence, and I don't think there is a need to repeat the discussion or to vote on the evidence.

CO-CHAIR SAKALA: Okay. If there are no objections, that will be the case. Great.

So, next is the importance to measure gap issues.

MEMBER GOYERT: Great. So, the developers have presented data from 2011 through 2014 that continue to demonstrate significant opportunity for improvement.

CO-CHAIR SAKALA: Okay. I think we don't need to vote on that. Is that your recommendation? Okay. That's right, we do because we are closer to where we want to be than we were before. Okay.

Any other comments on that portion?

(No response.)

Okay. So, can we open it up for

voting then?

MS. ROBINSON-ECTOR: Voting is now open for performance gap for Measure 1731. Option 1 is high; 2 is moderate; 3 is low, and 4 is insufficient.

(Voting.)

It looks like we are still missing four votes.

CO-CHAIR SAKALA: Are you counting one recusal that was added this morning?

MS. ROBINSON-ECTOR: Yes, I am.

CO-CHAIR SAKALA: Okay.

(Voting.)

MS. ROBINSON-ECTOR: Great. We've got 23 votes. All the votes are in.

Okay. So, 74 percent put it high; 22 percent put it moderate; 4 percent put it low, and zero put it insufficient. So, for the performance gap of the Measure 1731, the measure passes.

CO-CHAIR SAKALA: Okay. Do we have any new -- oh, Carolyn?

MEMBER WESTHOFF: Thank you.

Looking at the graph, I am just confused as to the change since 2011, and maybe somebody could explain that to me?

CO-CHAIR SAKALA: Not the voting, but the data.

(Laughter.)

MEMBER WESTHOFF: Great. It is just the data. It is not the voting. It is a question about the data.

MS. MILTON: It's too early in the morning. I'm still on Midwestern time.

I believe the reason that we are seeing a higher gap than what we did see is that there are more hospitals reporting, as a result of the fact that in 2014 we made it mandatory for hospitals with 1100 births or more annually to start reporting on the measure. So, I think that there weren't hospitals that were really monitoring this as closely. So, with a larger number reporting --

MEMBER WESTHOFF: Okay.

MS. MILTON: -- we are seeing more infections that we weren't before. Because we had less than 200 hospitals reporting when we went through this the last endorsement cycle.

MEMBER WESTHOFF: So, people are just paying attention now?

MS. MILTON: Correct, yes.

MEMBER WESTHOFF: Good. That's a good thing.

MS. MILTON: We're about 1200 hospitals now versus 200.

MEMBER WESTHOFF: Well, that's excellent, yes.

CO-CHAIR SAKALA: Okay. Thank you.

Do we have any new -- oh, Cindy?

MEMBER PELLEGRINI: Just a question. Can anyone explain why over 1,000 hospitals were added to this from 2013 to 2014? I'm sorry, I missed that.

MS. MILTON: We made it mandatory, if you had 1100 births or more annually and you are Joint Commission accredited, this is a

requirement that you had to report on the perinatal care set, all of the measures.

CO-CHAIR SAKALA: And beginning January of this year, it has gone to 300 or more.

So, do we have any new reliability data?

MEMBER GOYERT: The specifications have been updated to ICD-10. The numerator was changed a little bit to exclude babies that came septic the first 48 hours. The denominator was changed to remove the exclusion of greater than 120 days and bloodstream infection present on admission.

Developers reported an inter-rater reliability of 94-plus to 99 percent.

CO-CHAIR SAKALA: Thank you.

Is that a card up next to Janet? Karen? Okay, down.

Any other comments on the reliability issues?

(No response.)

So, let's have a vote on that because

of the changes, if there are no comments.

MS. ROBINSON-ECTOR: Voting is now open for reliability of Measure 1731. Option 1 is moderate; option 2 is low, and option 3 is insufficient.

(Voting.)

It looks like we are missing two votes. So, if everyone could re-vote, please? Thank you.

(Voting.)

So, 96 percent put it moderate; 4 percent voted low; zero voted insufficient. So, for reliability of Measure 1731, the measure passes.

CO-CHAIR SAKALA: Thank you.

Okay. Comments on validity, please?

MEMBER GOYERT: A priori, I don't think you would expect a correlation with the other perinatal core measure set, measures in the set. I think they the conclusions about quality can be made and that it is an indicator of quality.

CO-CHAIR SAKALA: Other comments on that?

(No response.)

Okay. Let's open the voting, then, for the validity.

MS. ROBINSON-ECTOR: Voting is now open for validity of Measure 1731. Option 1 is high; 2 is moderate; 3 is low, and 4 is insufficient.

(Voting.)

All the votes are in and voting is now closed.

Sixty-one percent voted high; 39 percent voted moderate; zero voted low, and zero voted insufficient. So, for validity of Measure 1731, the measure passes.

CO-CHAIR SAKALA: Thank you.

So now, feasibility, please.

MEMBER GOYERT: I don't think there's any issues with feasibility. Folks are doing it.

CO-CHAIR SAKALA: Other comments?

Yes, Sarah?

MEMBER NcNEIL: Sorry, this is delayed, but has there been talk about parsing it out by race as well for feasibility. If it is a required measure, is it possible to also add that?

Sorry. I was wondering -- this is late in the game -- but I was wondering if there is any way, also, if it is a very feasible measure, if we could also add something about race, because my understanding from the data is that race is not parsed out.

MS. MILTON: Yes, that could be done. I believe we did that for the SDS part, if I remember right. Yes.

MEMBER YOUNG: You actually did parse that out, and it showed that there was no meaningful gap in race.

MS. MILTON: Right.

MEMBER YOUNG: Which is different than many other measures we have considered where there is definitely a racial disparity.

CO-CHAIR SAKALA: Okay. If there are

no other comments, we can vote on feasibility.

MS. ROBINSON-ECTOR: Voting is now open for feasibility of Measure 1731. Option 1 is high; 2 is moderate; 3 is low, and 4 is insufficient.

(Voting.)

All the votes are in and voting is now closed.

Sixty-five percent voted high; 35 percent voted moderate; zero voted low, and zero voted insufficient. So, for feasibility of Measure 1731, the measure passes.

CO-CHAIR SAKALA: And finally, usability and use.

MEMBER GOYERT: I think it's fine. The measure has been improved with expanded data sources being available for use, removed the length of stay greater than 120 days, removed the bloodstream infection present on admission, and removed bloodstream infections that were not healthcare-related per se.

CO-CHAIR SAKALA: Do we have any other

comments on that?

DR. WINKLER: I have a question. I am just wondering, the changes to the specifications, what kind of impact is it having on the actual measure results, such to be able to monitor trends? You know, if you made the change this year, how are you looking at this year's data compared to last year's or two years' ago data? Is it making that a big of a difference in the measure results?

MS. MILTON: It should be more accurate. We made these changes based on feedback from the field. We were hearing from the field about -- the first part as far as bloodstream infections present upon admission has always been there, but when we took this measure and respecified it, we added this as a chart-abstracted data element because the present on admission indicator is not always there. So, we couldn't rely on just administrative data.

We allowed the hospitals to actually do the cases for babies that were born with

infections or were received in as transfers, and we did this by use of ICD codes, ICD-10 now, and looking to see if there were signs and symptoms, if they were on a longer course of IV antibiotics, to confirm that they were infected.

I am not exactly sure about a change in the numbers, though. The second data element hasn't been in use long enough for us to really look at any trends with the confirmed-after-48-hour data element.

CO-CHAIR SAKALA: Okay. Other comments about usability and use?

(No response.)

Okay. I think we can vote on this item.

MS. ROBINSON-ECTOR: Voting is now open for usability and use for Measure 1731. Option 1 is high; 2 is moderate; 3 is low, and 4 is insufficient.

(Voting.)

It looks like we are missing one vote to reach 23.

CO-CHAIR SAKALA: Do you want re-voting?

MS. ROBINSON-ECTOR: Yes, please.

CO-CHAIR SAKALA: Okay. Thank you.

(Voting.)

MS. ROBINSON-ECTOR: Great. We now have 23 votes and voting is now closed.

Sixty-five percent voted high; 35 percent voted moderate; zero voted low, and zero voted insufficient. So, for usability and use of Measure 1731, the measure passes.

CO-CHAIR SAKALA: Okay. So, we have determined that this measure again meets all the criteria that NQF has, and the final vote will be to decide whether we recommend it for continued endorsement.

Before we do have that vote, are there any other questions or comments?

(No response.)

Okay. Let's open it up for voting, then, for endorsement.

MS. ROBINSON-ECTOR: Voting is now

open.

CO-CHAIR SAKALA: Oh, sorry. John?

MEMBER KEATS: I'm sorry, I had a quick question which is kind of general. And it is probably because I am new. But I am looking at this. That improvement results in the bullet, it says, "In 2014, hospitals in the median lower quartile and 10th percentile recorded no infection," which is what you are looking for, right? What we are looking for is no infections.

So, shouldn't that be the 90th percentile, upper quarter, upper quartile? I mean, it seems like it is flipped. I mean, to me, 90th percentile means you are doing better than 90 percent of the hospitals out there. So, I am just wondering on the terminology. Is this typically how it is reported or should it be the other way around, or is it a distinction without a difference?

Okay. I am just used to seeing it the other way around. Then, I will shut up. Thank you.

CO-CHAIR SAKALA: Okay. Okay. So, let's open it up for voting if there are no other comments.

MS. ROBINSON-ECTOR: Voting is now open for recommendation for overall suitability of endorsement for Measure 1731. Option 1 is yes and option 2 is no.

(Voting.)

All the votes are in and voting is now closed.

A hundred percent voted yes and zero voted no. So, for recommendation for continued endorsement for Measure 1731, the measure passes.

CO-CHAIR SAKALA: Okay. So, there is another 100-percenter. That's great.

We have now to deal to deal with NQF's policy about related and competing measures, which is that we should have a thoughtful discussion about whether all three of the related measures are, indeed, needed; what the rationale would be; how we would turn around and justify this to the people out there in the field who are

feeling burdened by this, and how we would also justify our decisions to consumers and purchasers, payers, who are looking for changes in performance.

So, I think that we welcome anyone's comments, but the people who have taken the lead on these measures probably will have more detailed comments to make.

And there is a table that was provided by the NQF staff that puts all three measures side by side, and it is in the SharePoint. So, you can look at the comparisons.

MEMBER GOYERT: Could I ask the developers to go through the new document just briefly, comparing 0478 and 1731, that you provided just recently? I think that might help a little bit.

MS. MILTON: Our statistician did a comparison. He took data from The Joint Commission warehouse, and this would have been from the fourth quarter of 2014 to the third quarter of 2015.

Of that data, we looked at up to 30 ICD-9 diagnosis codes and 30 of the procedure codes. This is from administrative data, and comparing The Joint Commission measure to the other measure, which uses strictly administrative data. We tried to do sort of what we called an inter-rater reliability to see does the administrative data get the same results as if you do partial chart review, in addition to the administrative data.

The Joint Commission doesn't look at the present on admission flag. That is not something that we receive in the data, and that is something that the other measure, the NQI measure relies on in order to determine if those bloodstream infections were present upon admission.

So, what happened was we had 178,027 cases that were in the warehouse. Of these, 71 percent were missing principal diagnosis code, which the NQI 3 needs in order to exclude patients from the denominator population. So, we

knew that the coding wasn't always taking place as it should to get that initial patient population.

As a result of that, only about a third of the measure population was able to be identified using administrative data only. The fact that we were doing it in a little different manner allowed us to actually identify more cases that could be evaluated against the measure.

Then, the other thing that was a little bit different that might have attributed to part of the disagreement had to do with the fact that the NQI measure was allowing to exclude cases with seven days or less; whereas, The Joint Commission measure was looking at only two days or less for exclusions.

So, our statistician felt that it was about 10 percent of the discrepancy, right, as far as an exclusion. I understand that they have since changed that, I believe, to three days now versus the amount of time that we were looking at at that time.

Then, he ran a cross-table here and, of that, we determined that were 723 cases that did not appear in the numerator for the NQI measure, and approximately 94 percent of those had the secondary diagnosis code for the septicemia or bacteremia, but they didn't have the accompanying code for the staphylococcal or the gram-negative bacterial infection, which is required. In other words, if one is present, the other has to be present, in the NQI measure in order to be counted; whereas, we don't count both of those. So, we were able to identify more cases that way.

And the remaining 6 percent had to do with the fact that there was a diagnosis code for the staphylococcal or the gram-negative bacterial infection, but, then, they didn't have the newborn septicemia or bacteremia codes that would be present for your neonates 28 days of age and under.

Then, in addition, there were over 1100 numerator cases that were identified where

PC-04 identified those as being in the numerator, but the NQI 3 measure did not.

So, these were the significant differences, and it seems that administrative data doesn't always get everything because, with newborns, they don't always code as much as you would for an adult. So, that was where we felt this might have been part of the problem, is that the codes were not being uniformly assigned. Therefore, they are not able to identify cases just by relying on administrative data.

CO-CHAIR SAKALA: Thank you.

MEMBER YOUNG: Just a question or comment. Was this using ICD-10 codes or was this ICD-9 at 2014?

MS. MILTON: No, this was '09, yes.

MEMBER YOUNG: Okay.

MS. MILTON: We haven't gotten down that road yet, yes.

MEMBER YOUNG: So, these issues, obviously, will change as of this year with everyone using ICD-10 codes as a mandatory

reporting. So, I am curious to see what is going to happen in the next 12 months.

MS. MILTON: So are we.

(Laughter.)

DR. MAIN: For better or for worse.

MEMBER YOUNG: So, I think the side-by-side comparison or analysis is valid only for the 2013-2014 cycle, but I think, as of today, this analysis probably isn't going to hold true. It may, but with ICD-10 being very, very specific, I think that that is going to fall out, and I think you will see less discrepancy or less loss on the NQI 03, the 0478 measure we talked about yesterday and the measure we just discussed.

CO-CHAIR SAKALA: Thank you.

Before we get into our discussion, I would like to see if there is someone from AHRQ on the line who would like to comment on that analysis.

DR. OWENS: Thank you again very much.

This is Pam Owens. I am the

Scientific Lead of the AHRQ Quality Indicators.

Just a couple of points broadly, and then, I would like Corinna Haberland, who is the clinician on the Pediatric Quality Indicators, to discuss some of the differences because they are the ones that did all of the collaborative work with The Joint Commission and know the intricacies both of The Joint Commission's analysis as well as what AHRQ does.

I do want to point out that the version that was used here is not the version that is being released this summer. We, as a function of collaboration, changed things, and it was alluded to. The length of stay, for example, we have some additional qualifying codes, for example. But there is a version difference, that there would be more alignment that what is currently shown in the memo.

Corinna, are you on the line?

DR. HABERLAND: I am. Can you hear me?

CO-CHAIR SAKALA: Yes, we can hear

you.

DR. OWENS: Would you like to introduce yourself?

CO-CHAIR SAKALA: Can you speak up just a bit?

DR. OWENS: And if you could introduce yourself, that would give them some context.

DR. HABERLAND: Sure. I'm Corinna Haberland. I'm from Stanford, a pediatrician by training and one of the clinical leads on the AHRQ --

CO-CHAIR SAKALA: Could you speak up just a bit, please, Corinna?

DR. HABERLAND: Sure. Is that a little bit better?

CO-CHAIR SAKALA: Yes. Thank you.

DR. HABERLAND: Okay. Sorry.

So, my background is in pediatrics, if that is helpful to hear.

Just to speak a little bit about the harmonization that we tried to take care of, you know, the major difference between our two

measures seems to be or is obviously that The Joint Commission measure has the ability or allows the hospitals to go and do a chart review. Since our measure is based on administrative data, we do our best to be very clear about the infections that would be most likely to be perinatally-acquired, which is how we determined which codes to keep in the measure, and tried to also look at those that were more likely to be hospital-acquired.

I believe as Pam mentioned yesterday, obviously, we look at all causes of sepsis or all cases of sepsis as opposed to just those that are perhaps due to central lines, as The Joint Commission measure is focusing on.

CO-CHAIR SAKALA: Okay. Thank you.

Anything else from AHRQ right now?

DR. OWENS: The only other thing I would like to say is that we saw the memo yesterday. So, we haven't been able to do our own in-depth analysis about compare and assess in terms of numerator and denominator counts and

work with The Joint Commission specifically on this particular analysis. We have some hunches about the differences in the numbers, but without going through each one, that would be a little more difficult to do on the phone.

CO-CHAIR SAKALA: Right, and we also face the challenge that in both cases data are now different how they were before. So, that is a little bit of a challenge for us.

DR. OWENS: Exactly, yes.

CO-CHAIR SAKALA: I would like to share a little bit of the history of this. The originally-endorsed measure was the AHRQ measure. That was brought into The Joint Commission's core set. They turned it into a clinical data measure.

Last time around we did look at this same question of related and competing measures. There were Medicaid programs represented in the room who strongly said, "We cannot collect this unless it comes out of administrative data." So, that was the basis. That is why we have two

currently-endorsed, very similar measures.

And now, we will need to have a discussion today to decide how we want to handle that. And let's not forget about the third measure as well for -- what is it, lower birthweight babies? Very low birthweight babies. Okay, yes.

So, open it up for discussion, beginning with Tracy.

MEMBER FLANAGAN: You know, I think it is in the spirit of the users to want to do as much administratively as possible. So, my question is to The Joint Commission developers. With the advent of ICD-10 codes, will this be able to be primarily, with very little chart review burden, done administratively?

MS. MILTON: I think it is too soon to tell. It is really going depend on coding practices. And I can tell you we get a lot of questions about coding. Just even identifying the initial patient population for our maternity measures, we determined that hospitals weren't

coding for normal manually-assisted vaginal deliveries, which is a big thing that should be done, and it is not being done. So, it is really hard to say how this is going to translate into the infection measure until we have got some experience.

Coding guidelines and the coding clinics are trying to update the field, but it doesn't always get to the users. So, I think everybody is in a big learning curve right now.

MEMBER FLANAGAN: Let me ask a follow-up question. At present, what is the chart review burden on this measure? I mean just approximately. Is it like every case needs to be reviewed?

MS. MILTON: No. No.

MEMBER FLANAGAN: Just for those of us who may not be doing that primarily, does anybody in the room, can anybody comment on that? I know the maternity measures and what that burden is on chart review. I just don't know on the pediatric ones.

CO-CHAIR SAKALA: This is one, of course, of the maternity measure.

MEMBER FLANAGAN: I understand that, but I am an obstetrician/gynecologist. So, I actually review fallouts, and I know what happens at a very granular level. I'm asking the question from the standpoint of a system that adopts; what is the actual chart review burden of a clinician, a person, as opposed to administrative pull?

MS. MILTON: Our review would be those that have infection codes. So, the majority don't. It is pretty much an open-and-shut thing, especially if you are coding your birthweight, or you can pull a report up that can get birthweights. Then, there is really no reason to do a chart review.

But, if you do have one of the infection codes, that is what triggers the review. So, because the rate is about 3 percent, give or take -- depending on the hospital, it could be even less -- if there is an infection

code, though, then you do have to look at the record.

And I have talked to some of the larger systems, and it might be maybe 15 cases a quarter. It just depends on your volume of deliveries.

MEMBER FLANAGAN: Thank you.

DR. MAIN: I think the challenge here is that coding for neonatal sepsis is very complicated. Multiple layers of codes need to be put in to have it meet the AHRQ measure; whereas, this is really relying on using the codes for case identification and, then, chart review to make it more accurate.

MEMBER FLANAGAN: Thank you.

MS. MILTON: The other thing we learned, too, is that coders are coding if there is a lot of language in there about "suspected," "suspicious," and then, it never really gets confirmed. That was the other reason we were allowing hospitals to look at the record. Is this just because they pulled this code in

because it looked like they were septic or were they truly septic?

We are allowing them to look at blood culture results, how long they were on antibiotics, documentation of signs and symptoms. So, it becomes a little bit more objective, rather than we are just looking at a code that said they had an infection, when, indeed, they didn't.

CO-CHAIR SAKALA: Thank you.

Sindhu?

MEMBER SRINIVAS: So, is the purpose of the chart review, then, to sort of eliminate false-positive cases, and so, to make it more accurate? Or is it to get people to think about sort of the -- like I guess it is sort of a combination of viewing the cases and trying to get at what could be improved to reduce sepsis personally.

But, from a like data perspective, what is the percentage of like false-positive identifications? So, the idea that in order to

do chart review, people want to look better, so that they can actually eliminate some of the cases that seem to have been identified, but, then, are misidentified. But it sounds like the volume of that is actually pretty low. And has it been worth the chart review to eliminate a couple of pieces? Or do you just take a little bit of noise to make the measure a lot easier from a system perspective?

MS. MILTON: Well, we have a lot of type A hospitals. They would prefer not to have an infection, if it truly wasn't, you know, show up on their record. So, this is why we made that part of the measure, a chart-extracted, to identify for sure were they infected or not.

So, like again, we are only looking at cases -- they could have a thousand discharges, and there might be, like I say, 15 that have infection codes. And then, they are incumbent to look at the record to determine was it, indeed, an infection or just that it was a suspected infection that got coded. That seems to happen

more than not.

DR. MAIN: Sindhu, on the other side, though, I think there was a pretty high rate of false-negatives. There was 33 percent of the cases weren't identified with the AHRQ measure. That was picked up this way by using what are the codes and, then, doing the chart review to confirm.

CO-CHAIR SAKALA: Greg?

MEMBER GOYERT: Maybe somebody can help me. I just don't understand, if all the hospitals are collecting the measure for The Joint Commission, where does the disconnect come in with Medicaid? They are doing the work. They are collecting the data. Why can't Medicaid see the results? Am I being slow?

CO-CHAIR SAKALA: Can someone comment on that in the room?

MEMBER SHEA: I think the only distinction is that, from a Medicaid perspective, when you are looking at claims, you can have a three-month lag data showing you what your rates

are; whereas, with The Joint Commission rates, we really have to wait perhaps until the end of the year, a little bit longer, before the results are published.

DR. MAIN: The bigger issue is identification of Medicaid enrollees. Medicaid wants their numbers based on their enrollees, as opposed to The Joint Commission that does it for all the patients in the hospital.

MEMBER SHEA: I'm sorry, I didn't really catch onto that, but it is true. So, we would receive all-payer data from The Joint Commission; whereas, I could look at Anthem's rates for the facilities where our patients are receiving care.

CO-CHAIR SAKALA: Okay. Raj?

MEMBER WADHAWAN: While the chart here seemed like an inordinate amount of burden, the reality is that almost every institution gets a trigger off records, considers those. And all those cases are reviewed. Because all infection teams that I know of, almost every hospital

reviews all those cases because it is important for them internally.

So, this data exists. It has already been done. They have gone through the charts. I don't think it is that much of an inordinate burden that these additions have to be. It is more accurate.

CO-CHAIR SAKALA: Thank you.

Jaleel?

MEMBER MAMBARAMBATH: Yes, I agree with Raj; it is not an inordinate burden. We do this on a regular, routine basis at our institution.

As you mentioned, the incidence of bloodstream infections is very low. And so, every month when you review this, there is either zero or one case, even though we have about 1,000-1200 admissions a year. So, yes, the burden is not inordinate; I would agree with that.

One of the things I would like to comment is, with the AHRQ measure, this is blood

culture positive, all-comers. So, even if they are receiving antibiotics for necrotizing enterocolitis as urosepsis pneumonia, but if they have a blood culture positive, then they are included in the group.

But The Joint Commission measure does not have that. That excludes all the other patients. It is only strictly bloodstream infection. And if they have another diagnosis, they are excluded from the group. Is that correct?

MS. MILTON: That's correct. And that, again, was based on feedback from the field.

MEMBER MAMBARAMBATH: Yes. This kind of leads to gaming of the system a little bit because none of the other things, necrotizing enterocolitis, for example, or pneumonia, are public-reportable issues. And the diagnosis of necrotizing enterocolitis, stage 1, or pneumonia is very debatable sometimes and it is very subjective. So, people can add in the diagnosis

of necrotizing enterocolitis for a suspect necrotizing enterocolitis or pneumonia to reduce their number of bloodstream infections.

MEMBER YOUNG: Jaleel, on the new ICD-10 codes, you can't put "suspect" anymore. It is actually falling out in the nomenclature, at least from coding at least in the emergency department. We don't have the ability to have "suspect" or "probable" anymore. That is gone. You either have it or you don't. You can put a comment; there is no way to code for it.

MEMBER MAMBARAMBATH: Yes, but necrotizing enterocolitis is different because -- yes, it is on. It is on? Is it okay now? Yes.

So, with necrotizing enterocolitis, the situation is different because stage 1 necrotizing enterocolitis is suspected necrotizing enterocolitis.

CO-CHAIR SAKALA: Tracy, did you want to make a comment?

Okay. Sindhu?

MEMBER SRINIVAS: I guess I am asking

a question about sort of how do we -- like what is the goal of sort of the discussion in the sense that -- I mean, it seems like the fundamental is here, but the chart burden is not too difficult, and people are doing it anyway, which is what you would expect if there is bloodstream infection.

It seems like we are more like debating sort of the coding itself and the exclusions and the inclusions. But it seems like there shouldn't be two different measures that are essentially measuring the same thing.

As people move forward with quality, I feel like one of the problematic things is that there are sometimes too many definitions for the same thing, and then, everybody ends up using a different definition. And then, you can't compare the apples and oranges. And so, kind of having a consensus on a single way to measure certain things seems very important. But I guess I don't know how what we are sort of charged with.

DR. WINKLER: Actually, you are charged with grappling with that question, which is not an easy one, absolutely. And so, you could call this the fifth criteria. All right?

So, we have looked at the measures fundamentally. If they had failed on any of the others, they wouldn't be on the side-by-side. So, the question on the table is the one you raise. You know, what are the issues around having multiple measures? These three measures come from different data sources. Their history and how they got here are different and have various reasons for doing so. But, on the other hand, as you say, there is a lot of feedback and pushback about like, why do we have so many? Why don't we focus on one?

And that is really the question on the table, is: do you want to make a recommendation around that? You know, can you justify saying, "Okay, well, three is fine because...." or maybe having three isn't so fine and we would recommend going with one, two, whatever. And so, that is

what your conversation is about. It absolutely is a difficult one to have, but that is exactly your charge at the moment.

CO-CHAIR SAKALA: Yes?

MEMBER YOUNG: Sorry. Is there anyone from CMS or Medicare/Medicaid here to discuss? Because, apparently, they were the largest stewards of NQI 3.

MEMBER JOLLES: In response to several of the previous comments, beginning with Dr. Keats' comment about the performance improvement opportunities, at some point we have to talk about when it is time to retire this. So, strategically, the function of NQF is to really push the needle and improve quality, and it is not in our best interest to continue to endorse measures where we have achieved the goals.

So, with that in mind with regard to this debate we are having, I would say that the next step in strategy in data measurement is to adopt this proposed measure that includes chart review, because what you need to do now is

hardwire. So, if we have figured out how to fix it, we have almost got it down to zero harm, now hardwiring process is where people are routinely auditing their codes and manually reviewing the charts and having meaningful meetings and processes to improve, to maintain and spread the practices.

But, at some point, we need to have a bigger discussion about the entire profile of what we are endorsing and the fact that we have an emphasis on acute care. From a population standpoint, we need to be looking at population health when we put forth a panel related to perinatal and gynecologic care. I know that is outside of this.

CO-CHAIR SAKALA: Okay. Thank you.

Matt?

MEMBER AUSTIN: Yes, just to reflect, I think, what others have maybe referenced earlier, to me, I think one of the big wildcards here is ICD-10 and us not really knowing or understanding how well ICD-10 will capture some

of the maybe deficiencies in the AHRQ measure.

So, I guess, to me, one option would be, are we allowed to sort of kick the can down the road another two or three years to see how ICD-10 plays out and maybe have more information to make a better decision?

I mean, in my mind, the big concern with chart obstruction is burden, and where we can reduce burden or eliminate burden, those should be opportunities that we take advantage of. But I think ICD-10 is sort of why we put it here, yes.

DR. WINKLER: The answer to your question simply is, yes, you can.

CO-CHAIR SAKALA: Amy?

MEMBER BELL: I like your point about making sure that we don't just keep endorsing things just to endorse them. And I also like Matt's point about ICD-10. But I think we need to bring about, as we bring in the facilities that have 300 deliveries or more, there is a whole gap of hospitals that we truly don't know

their performance yet, and we really need to make sure those processes are onboard and we see the results of that before we just stop endorsing this one.

CO-CHAIR SAKALA: Thank you.

Other comments?

MEMBER NEVINS: I have just a question. Does anyone have any statistical data or some idea of the extent of the use of electronic medical record in the nation? Because I think at some point, you know, the burden of data extraction will be not an issue, right? Now, yes, electronic medical records also means that you get 300 pages, right, that someone has to look through, but I just wanted to sort of introduce that as something that we should consider when we are thinking in terms of burden.

CO-CHAIR SAKALA: Do you want to comment on your plans for eMeasures or anything else?

MS. WATT: Hi. I'm Ann Watt from The Joint Commission.

CO-CHAIR SAKALA: Can you come a little closer?

MS. WATT: Sorry. Sorry, Elliott.

(Laughter.)

I don't have specific numbers. I can tell you, though, as we all know, the uptake of electronic health records is becoming significant in acute care hospitals. And we are quite active in the development of electronic clinical quality measures which are extracted directly from the electronic health record. Of course, that is everybody's goal, that there will no longer be a burden of data collection. I think we are a way away from there, to be perfectly honest with you, but we are getting there.

The challenge is that, even with the electronic health record, the data element that we are looking for needs to be in a structured field. Nobody is going to be reviewing an electronic health record to pick out that nugget of information that we need. That is the challenge.

But, you know, measure developers, not just The Joint Commission, all of us have been working very hard to develop that capacity and, also, to work with the EHR vendors, so that it is easier to do.

CO-CHAIR SAKALA: Raj?

MEMBER WADHAWAN: I just wanted to add to that, I think it is about 80 percent, that option of acute care hospitals today from the EHR perspective. I just read it a few days ago; 78.5 percent was the number that I came across.

While it seems like it should be a very achievable goal of pulling data out of the EMR, those of us who have been working in that space for the last many years, the databases, we have been utterly unsuccessful, for a variety of reasons.

You know, Virginia and I both participate several databases for neonatology, it still is all manual extraction. I don't think it will be achievable in our time on the Committee for sure; I am hoping it will be achievable in

our lifetimes.

CO-CHAIR SAKALA: Greg?

MEMBER GOYERT: So, I guess I am wondering, what are the choices before us? We are not going to un-endorse a measure that we endorsed yesterday or what, I guess, for the Committee? Door A, Door B, Door C?

DR. WINKLER: Because you feel all three of them do meet the criteria, those recommendations stand. This is the next step. It is, of those, do you feel that you want to say, because we don't want to have three doing the same thing, we want to go with one, or two, or whatever? If there is enough reason to say it is okay to have three continuing, that is an option as well.

And so, that is why having this conversation that we have been having is important to factor in all those considerations. At this point, we probably want to hear from you, what would be your proposal among all those options?

CO-CHAIR SAKALA: Kim?

CO-CHAIR GREGORY: I wanted to follow up on that EMR issue, just to say it is hard to get the data out. You almost have to have a full-time person who works for both the IT Department and your department because they don't always talk to each other, and you are really a low man on the totem pole if you are not in the IT Department.

And the other issue is that a coder can only code what a doctor writes as a diagnosis, not a nurse, not anybody else. And so, what we are finding is that the doctors aren't writing enough diagnoses. And so, even though with ICD-10 we have the capability to be very specific, you know, you just get tired of clicking and you just put down two diagnoses and you move on to the next patient. So, I am concerned that the diagnoses won't be in the chart. I think that we just need to put that out there.

And then, just sort of as a

consideration for these three measures, I think probably the discussants and the clinicians who are most involved in it, if they wanted to propose something, we would certainly be interested in what you thought. And then, maybe another idea might be if we want to prioritize or rank them in some way.

CO-CHAIR SAKALA: Sheila?

MEMBER OWENS-COLLINS: Okay. So, I am also just trying to make sure I understand what we are looking at.

All right. I just wanted to make sure I understand what we are looking at. It looks like the emphasis is primarily on low birthweight infants, and I understand that, and only the PC-04 could potentially look at low birthweight as well as the larger babies, the term babies, with any incidence of infection. Is that correct?

DR. WINKLER: Yes.

MEMBER OWENS-COLLINS: Okay. So, I mean, I think if we are going to look at

infection in newborns, we should look at all of them. So, that is my vote. Then, that would mean we would have to look at least two out of the three.

DR. HABERLAND: Corinna Haberland again from Stanford.

I just wanted to add the AHRQ measure also looked at larger babies, if they had a major surgery or were ventilated.

CO-CHAIR SAKALA: Could you turn your microphone on and say that, please, Jaleel?

MEMBER MAMBARAMBATH: Yes. So does The Joint Commission measure, too. So, if they have more than 1500 grams, but they are mechanically ventilated or have a procedure or die, The Joint Commission measure includes those patients as well.

CO-CHAIR SAKALA: All right. Did you have something else to say, Jaleel?

MEMBER MAMBARAMBATH: Yes.

CO-CHAIR SAKALA: Yes.

MEMBER MAMBARAMBATH: So, we have

discussed these two measures at length, but we have not talked about the Vermont Oxford Network measure. I wanted to consider that as well.

One of the things which the Vermont Oxford Network measure does not include is those babies who are more than 1500 grams. Also, it includes meningitis in its mix. Now early-onset sepsis, the incidence of meningitis is very low, but with late-onset sepsis, the incidence of meningitis is slightly higher.

But most of these cases of meningitis are secondary to hematogenous spread from the blood culture. So, the blood culture is positive. There is a small, little group of babies wherein the blood culture will be negative, but the CSF culture can be positive.

CO-CHAIR SAKALA: Small like 1 percent or 10 percent?

MEMBER MAMBARAMBATH: So, that is the question I had for the Vermont Oxford Network measure yesterday, whether we have an idea of how small is this group of meningitis babies. I am

not sure whether we are capturing a significant number of babies in that group of meningitis and, if not, then is there a reason to continue with that measure while we are capturing this information with The Joint Commission measure or the AHRQ measure?

CO-CHAIR SAKALA: Is anyone from VON here to comment on that?

DR. WINKLER: She is right behind us here. She is here, I think.

DR. EDWARDS: So, as I explained yesterday, there is no way to tease out the meningitis in the definition.

MEMBER WADHAWAN: Why can't you do that with that data? Based on my knowledge of VON, which is a lot more limited than yours, if you have babies coded as sepsis and you have got babies coded as meningitis, you can look at the two together and see where they don't disagree. Where you have got meningitis that is CSF-positive and you don't have sepsis, that is your category.

Now, if you go back and look at the published data from NICHD, it is about 10 to 20 percent of ELBW, not VLBW, extremely low birthweight infants, who would have CSF positivity without blood culture positivity.

But Jaleel's question is very interesting and relevant, I think, to more sort of to this Committee, as to what is the incidence. And you have a database that can answer that question.

DR. EDWARDS: But the definition that we give is that you can have a positive culture from blood or CSF. So, it is not two definitions; you have a positive blood culture; you have a positive CSF culture. It is blood or CSF. And then, it is combined with coagulase-negative staph.

MEMBER WADHAWAN: You don't collect it separately?

DR. EDWARDS: No. No, but, I mean, this is a fantastic question. In a perfect world we would ask our members to separate those out

and report them to us, and maybe we should, as a result of this conversation, because it certainly sounds like it would add to the information that we have.

MEMBER WADHAWAN: Can I just add one more thing, since we talked about the VON measure? I think the one thing that is different about this measure is that it has a risk adjustment, and that is kind of relevant if you are, as a consumer, looking at information.

Because, you know, if you have a nursery with a small Level 2 and you have got no healthcare-associated infection because you know you have to put in a central line and you transferred all those kids, versus a large Level 4 NICU in the city that is taking care of the sickest patients, the numbers can be very different and it could be misinterpreted in a wrong way in a hospital that takes care of the sickest of the sick kids. And I am sitting in a Level 1 nursery; I have got zero AHIA and he has got 1.2 percent. It doesn't necessarily mean he

is bad.

What VON data does is that it allows you to make that comparison because you have other factors. As imperfect as the model is, at least it is an attempt at answering that question.

CO-CHAIR SAKALA: Thank you.

Yes, Janet?

MEMBER YOUNG: I want to just frame the discussion slightly differently, looking at it as these very small or smaller hospitals, the 300 deliveries, are just now coming online for The Joint Commission's measure. Many of those hospitals don't have access to electronic health records, and they are not even automated in any way, shape, or form. They are still back in progress notes and handwriting and hand-signing everything.

So, I would like us to at least consider those small hospitals who now have a huge burden to report out this data to The Joint Commission and perhaps give them a learning

curve. At least if not for a year or two years, when we next discuss this measure, to look at the differences in those much smaller hospitals than what we currently have reported right now.

DR. MAIN: Although it should be pointed out that those hospitals aren't going to be caring for VLBWs or babies on ventilators. They will be transferred to other settings.

MEMBER YOUNG: Oh, that is a good point.

CO-CHAIR SAKALA: So, Nancy and Jennifer, I don't know who was up first with your card.

MEMBER LOWE: Yes, as a non-neo, non-peds person in this group, I am struggling with -- I need some help from the experts on the Committee to know which of these measures moves us more consistently toward quality care, which is what this is about. So, I would really like some clear opinion from the peds/neo people about which one, as you look at them now, you think helps us accomplish that goal better.

MEMBER BAILIT: Can I just make mine before you answer that, because I think it dovetails onto that?

CO-CHAIR SAKALA: Great.

MEMBER BAILIT: I absolutely agree with your point. I was going to make something very similar. The question is, you said a keyword. What is your opinion about which one is? We don't know, right?

In the ideal world, you would collect both for two years. You would compare which one is directionally better. Are they going in the same direction? If so, it doesn't matter; let people pick.

So, to the extent that this is an opinion, let's be clear, it is an expert opinion. If we can convince our users of these kinds of measures to collect both and get some empirical data about which is directional more helpful, that to me sounds like a good solution.

CO-CHAIR SAKALA: Great.

MEMBER WADHAWAN: Can I take that

also?

CO-CHAIR SAKALA: Thank you.

MEMBER WADHAWAN: It depends on what patient group you care about more. There's about 40,000 very low birthweight infants born in the country every year. If that is the group that we cared about, which we should because that is where the bloodstream infections are most prevalent, then, of course, the VLBW measure makes more sense because it is more specific, it is more accurate. But, then, there's also those other newborns that have low sepsis, but there's a lot of them. You get either a postnatal onset of a disease that would cause you to be in a NICU or you have congenital malformation. Those are bigger kids that are not captured if you just look at the one, one specific measure. So, that is the problem. They answer different questions.

I think if you just focus on the bread-and-butter neonatology, which is VLBW care, what moves the needle is, in my opinion, the one measure. But, then, there the whole stuff, but

most NICUs don't care for those patients. It is about a fourth of the NICUs in the country that care for a lot of those patients. Three-fourths of them don't. Three-fourths of NICUs just care for regular prematurity and VLBW infants.

CO-CHAIR SAKALA: Thank you.

Are your cards up? Yes?

MEMBER MAMBARAMBATH: So, I agree with Raj's comments, but I think one of the other issues is that what Vermont Oxford Network will measure is not publicly-reported. It needs a registration fee, while the other two measures are publicly-reported.

CO-CHAIR SAKALA: Thank you.

Sindhu?

MEMBER SRINIVAS: I guess I am struggling a little bit with the sort of focus on the population, just because, one, as you said, limiting something to a population or a network that not everyone belongs to might not be the right solution. But, separately, bloodstream infection prevention is like process metric,

right? I mean, it is an outcome that we are preventing, but with a known process to prevent it.

And so, when you focus on certain small babies, even though they have a higher prevalence, you are missing the opportunity to engage other hospitals in a process improvement structure or having a standard process, because we know standard processes, when they are implemented, can reduce central line infections.

And so, it seems like, while smaller babies have a higher prevalence, that focusing on all babies and just trying to prevent all bloodstream-associated infections in general, at least from my perspective, would be where we should focus.

CO-CHAIR SAKALA: Yes, Sheila?

MEMBER OWENS-COLLINS: Okay. I agree with what everybody is saying. I think that we should have a more comprehensive approach, and bloodstream infections are a very serious issue. Also, if we concentrate on all newborns, it will

dovetail to the efforts in the adult world, preventing infection, you know, most nosocomial infections.

And to the point of the community hospitals, I think that, even though the numbers are small, they would have a significant contribution in terms of looking at infections in the larger babies, because, as someone said, there are more of them and they can have serious morbidity from those infection.

CO-CHAIR SAKALA: Thank you.

So, that really raises the question of, if we do both parts, what about the two measures that are quite similar as well?

We need to tie this up. Maybe the answer is that we can't do better than we did before last time around.

Also, I think that the points about the ICD-10 and the changes in the AHRQ measure and the broader pickup and ongoing look from The Joint Commission are all good reasons to say let's take a look again, if we can't make a

decision now.

So, I would love a proposal from anyone about how we should move forward.

MEMBER BAILIT: So, I guess what I would say is we have approved all three. People who are using these measures are going to use them for different things. But, to the extent that we ask our users, if you will, to collect data or, if possible, to collect both sets of data, so we can have a comparison for next time, but I think they are being used for different things.

You know, the NICUs who are the Parklands and the other high-elite taking care of very small babies, most of them are Vermont Oxford anyway. The general hospitals are going to use The Joint Commission measures. So, to the extent that those are being collected in both places and lots of hospitals, let's see if we can get that data from them.

CO-CHAIR SAKALA: Thank you.

Karen?

MEMBER SHEA: I agree. Practically thinking about someone who is sitting in maybe a general hospital with the burden of collecting information for The Joint Commission measure and, then, a separate data collection that would, then, be sent off to VON, and then, maybe an insurer coming in behind both of these and saying, "Geez, I really want you to collect information on the AHRQ measure because this is what we can collect in terms of administrative data," it can be a burden on those health facilities if they have to collect data on these three measures from a practical standpoint.

But I do see that there is merit in each one of these different measures. And I guess I won't repeat what the group has already said, that it might be worthwhile to see what are the distinctions between the three.

But I guess, in summing up, I would say to those facilities who are being asked to collect data on all these three measures, if they can push back on that request and select the one

that seems most suitable for their institution.

MEMBER BAILIT: Having said that, we know that some hospitals are already collecting all three. So, rather than asking them to do it de novo, we can just say, "Hey, if you happen to be doing more than one of these, let us know."

DR. EDWARDS: So, that is actually a great point, and we do a member survey every year. I'm sure The Joint Commission does as well. I will add that to the member survey -- I am meeting about that tomorrow -- to find out what other hospitals are reporting on these measures and how they are using them.

And I would also like to say that I would love to work with The Joint Commission to compare data as well in aggregate, if possible. So, that might be something that we work on between now and then.

CO-CHAIR SAKALA: Naomi?

MEMBER SCHAPIRO: Yes, this is just a question about unintended consequences, since I am new to this process. Which is, if we sort of

continue to endorse all three, then we don't necessarily have control over who requires people to collect them. And I am just kind of wondering if we have any idea of what that burden might be, especially for the new hospitals coming online that have fewer resources.

You're shaking your head. So, we don't have any idea?

DR. WINKLER: I mean, you all are out in the field and can give us a much better idea of what is going on in the real world.

MEMBER YOUNG: What was the question? Who was reporting to whom?

MEMBER SCHAPIRO: No, the question was, if we continue to endorse all three and say this new round of hospitals that are coming into reporting, are they going to have to report all three? Or they don't report the VON because they are not in the VON?

MEMBER YOUNG: So, the subspecialty, the high-level NICUs are reporting to VON. They are all part of the VON network. Almost everyone

has to report to The Joint Commission, again, if you have 300 deliveries or over. So, that is essentially every community hospital and higher.

And I'm not really sure who is reporting to the Agency for Healthcare, or AHRQ. I don't have a sense of that.

DR. OWENS: On the AHRQ measure, just so you know, there is no reporting burden from a hospital perspective. It is coming directly out of billing data. The hospitals collect the discharge data. And then, for what we use it, each is sent to their state data organization, rolled up and collected into a graphic state.

For claim status or if it goes to Medicaid, for instance, the state Medicaid offices get the data, but the hospitals don't have to do anything to subset it. I mean, it is sent as part of the bill.

And by the way, AHRQ does not do hospital-level reporting.

CO-CHAIR SAKALA: We have many more comments, and we really need to wrap this up in

the next five or so minutes.

So, the comments that driving us toward a decision --

MEMBER KILDAY: I just have one clarifying question. How many hospitals in the country are Joint Commission versus D&B? D&B. There are two different accreditation agencies out there, and we can't make the assumption that all hospitals are reported to The Joint Commission because that is not true.

MS. WATT: Ninety-some percent of the hospitals are Joint Commission, though, 90-some percent.

CO-CHAIR SAKALA: And about 80 percent, a little over I think, are now included in the 300 or more births.

Sindhu?

MEMBER SRINIVAS: I guess I would like to -- and maybe this is just sort of striving for the future -- I mean, it just seems like the only sort of thing that I heard about the AHRQ measure was the payer issue of payers getting, I know it

is like direct data feeds to HCUP and payers are being able -- like you can do a payer-related evaluation instead of getting all payer data.

But it seems like fundamentally the difference is really in the codes and the exclusion. Because you could take The Joint Commission measure and the codes and make it payer-specific. So, that doesn't seem to be the limitation.

It is seems to be fundamentally in terms of what is actually being included and excluded. And I may be misunderstanding that, but it just seems like I would caution us, for the reasons that Naomi mentioned, too, of having so many measures that are really trying to get at the same thing and trying to push us towards harmonization.

I mean, to the public, we are all the experts in these areas and we can tell like, well, in this measure, these couple of people are excluded because of these codes, and in this other measure they are not. But when you put

them side by side and the numbers are so different, from a public perspective, that is like incredibly confusing. And it seems like we should be striving towards a single harmonized way to measure things to really drive quality.

I feel like it is very difficult to have different ways to measure the same exact thing that are marginally different. And I think we should be trying to get to a place where we are actually limiting it from three to at least two, and even getting towards one eventually.

CO-CHAIR SAKALA: Ann?

MS. WATT: Thanks.

One thing I want to emphasize is that Joint Commission staff and AHRQ staff have worked very closely together to harmonize all of the data elements that are possible to harmonize and all of the specifications in these two measures.

I think what you are seeing as the fundamental difference between The Joint Commission measure and the AHRQ measure is the dependence solely on coded data used for the bill

in the AHRQ measure. There is no reliability evaluation of the coding, of the appropriateness of it. There is for The Joint Commission chart-abstracted measures. It is one of our requirements, is that inter-rater reliability be done, so that we have a fairly -- and we do extensive quality testing of the data. I think that is one of the things that sets our measure apart.

Having said that, though, I think that The Joint Commission measure -- I think that what our analysis that Celeste discussed at the beginning of this discussion showed was that the biggest differences is in the coding. And so, I think that is a big difference. It is just sort of a fundamental thing.

We haven't done a comparison with the VON data. We would be happy to work with VON, so that we can do that type of harmonization as well.

I don't think that we are talking about so much as apples and oranges. We are

talking about different data collection systems, basically.

CO-CHAIR SAKALA: Sheila?

MEMBER OWENS-COLLINS: I just wanted to just make another quick plea for comprehensive data collection and looking at the late preterm and the bigger babies, because, as I was noticing the criteria for infection, I noticed that GBS, unless I missed it, is not up there, which is a real significant marker for infection for bigger babies. And it is a very prevalent, probably still the most common infection in newborns.

CO-CHAIR SAKALA: Jaleel?

MEMBER MAMBARAMBATH: So, it appears that we are moving forward endorsing all these three measures. I was on the Committee three years or four years ago discussing the same thing. It was a robust discussion, but ended up kicking the can three or four years forward.

(Laughter.)

And now, we are kicking the can again for another three or four years. So, I am kind

of feeling disappointment at that and trying to see if there is a way to try to hone-in on this and try to fix this problem before the next time the Committee meets.

I think it was really helpful to see the comparison between the AHRQ measure and The Joint Commission measure. Thanks a lot for that.

But I think it would be worthwhile trying to see if we can have this comparison in more detail, looking at meningitis, and giving a chance for AHRQ to come up with their opinion on this comparison, and all that.

I am trying to see if we can have this as one of the items even for the Committee to start off with, and even before we reach this level, get a better understanding of the lay of the land.

CO-CHAIR SAKALA: Thank you.

So, that is a good way for us to turn to the developers and say, "Help us next time around with better working together around moving toward a better place."

So, Diana and Nancy.

You know, Jaleel is saying we are moving toward continuing with all three.

MEMBER JOLLES: Could we vote, because I don't support that? And I don't know; maybe I am an outlier.

CO-CHAIR SAKALA: Do we have to vote on, first of all, all three and, then, if we don't support all three --

DR. WINKLER: Vote on the proposal.

CO-CHAIR SAKALA: The proposal is?

DR. WINKLER: The proposal is to keep all three.

CO-CHAIR SAKALA: To keep all three?

MEMBER MAMBARAMBATH: If we say we vote for the proposal and the answer is no, then what do we do?

MEMBER JOLLES: Then, we call on the neonatologists to help us.

(Laughter.)

CO-CHAIR SAKALA: Nancy, while we are waiting.

MEMBER JOLLES: Well, I mean, I am half-tempted to begin like a lecture on what NQF is and what we are being called to do. We can't kick this down the road.

MEMBER LOWE: I agree with that.

DR. WINKLER: Let me offer one other potential, because I think a great many of you have suggested that there is a need for better data on which to make your decisions.

As to opposed to waiting, say, three years or so, this is a standing Committee. We could perhaps consider something in the interim before three years to bring the data back, if our developers are willing to work with us on that. So, we could have this discussion with a little bit more data to help us understand what is going on. So, that is potentially, again, one of the benefits of the standing Committee, sort of an outstanding issue that we might be able to deal with on an off-cycle basis.

MEMBER LOWE: Carol, actually, Reva just said I was going to say, because that is a

difference now that we are a standing Committee, that we are here for a period of time. And why can't we ask them to come back to us in a year with more information about the pros and the cons, what the data look like, using the two different measures?

I totally agree with Diana. I think we are not helping the industry from the perspective of what NQF is about, I don't think, by kicking the can one more time.

MEMBER FLANAGAN: I just want to make one additional comment that was said earlier, but to emphasize. What I heard the table say is that the AHRQ measure is being used by Medicaid, primarily because it has almost no -- it is derived administratively.

To the extent that PC-04 can be made more administrative, I think it would satisfy Medicaid. And I think we need to ask Medicaid whether they would consider changing. Because, you know, they, then, accept PC-04 with very little chart review burden. This whole

discussion goes away.

CO-CHAIR SAKALA: So, are you okay, Jaleel, with a proposal to stand with the three measures until next year, until a year from now, when we get more information from the three developers, including regarding the changes that have been made?

What is the proposal that we are talking about here?

MEMBER MAMBARAMBATH: I agree, but why is it only me?

(Laughter.)

CO-CHAIR SAKALA: Well, you were the person who opened up saying maybe we accept all three. So, there we have it.

Nancy?

MEMBER LOWE: Yes, can we ask the three developers, charge them with coming up with a true single measure? Why can't we charge them with that? Why do we have to figure that out when they are the people that work with this data all the time?

CO-CHAIR SAKALA: Well, I think VON has its own limitations around the population-wide measure.

MEMBER LOWE: Well, that's their problem, not ours. I'm sorry, but it is. Because if we are all interested in quality care, then it seems to me that we should be able to negotiate a single measure that provides useful data for this issue of infection in the high-risk neonate. I'm sorry if I'm too simplistic, but --

CO-CHAIR SAKALA: Matt?

MEMBER AUSTIN: Yes. So, my understanding is the VON measure is actually looking at a subset of the babies that the other two measures are looking at. The VON measure is looking at the really small babies.

So, from my perspective, it feels like I could understand why a hospital might want to have specific data on those. To me, the real conflict is between The Joint Commission measure and the AHRQ, as they really are looking at the same population and trying to get to the same

outcome.

I do feel like we need additional data with the change of ICD-10 to really make a distinction on which we think is the better measure. So, my proposal would be we reconvene in 18 months to evaluate those two measures and, for now, we stand with the three measures.

MEMBER YOUNG: I would like to second that proposal, so we can move on.

CO-CHAIR SAKALA: Raj?

MEMBER WADHAWAN: I just wanted to add to that comment. I think the only measure, if we had to choose one of the two, it cannot be the VON measure because it is not all-encompassing and it doesn't answer the questions for other babies. So, it has to be one of the two, although VON, certainly being a subset, with some work could be incorporated into that.

CO-CHAIR SAKALA: So, a stratified measure that VON would be able to use a part of?

MEMBER OWENS-COLLINS: Yes. And again, I mean, if we are going to come back, I

will make the plea again that we look at the bigger babies and not just concentrate all of our efforts on the smaller babies, as important as they are.

MEMBER BAILIT: I think the other thing here, though, is that VON is voluntary. Nobody is going to force you to do VON, right? So, you are really just talking about, if you are worried about data burden, it is just the other two.

DR. EDWARDS: It is, and from VON's perspective, we are going to collect it anyway, and we are going to collect it for the expanded centers on all infants admitted to a NICU, which in hindsight maybe I should have proposed that measure as a new measure.

I mean, I would sort of agree with Matt, but our measure is different in that way, in that it is the very low birthweight infants risk adjusted, SMR or 0 minus e, but including meningitis. So, it is kind of fundamentally different from the other two.

But, as I said, I am very happy to work with AHRQ and The Joint Commission on harmonizing.

CO-CHAIR SAKALA: We need to move on. So, can we take a vote on a proposal?

Matt, would you like to clarify? Yes.

MEMBER AUSTIN: So, the proposal is that we reconvene in 18 months to evaluate the AHRQ measure and The Joint Commission measure based on updated data. Yes, the VON contribution as well. So, I guess the proposal is in 18 months to reconvene with better data to continue this conversation.

CO-CHAIR SAKALA: And put some pressure on the developers to help us move toward a single measure.

CO-CHAIR GREGORY: So, to clarify that, on the VON piece, though, since that is chart audit, it would be worthwhile to know the overlap of their chart audit with administrative codes. How many of those babies would you get? That is what we are asking.

CO-CHAIR SAKALA: Okay.

MEMBER WADHAWAN: Just a quick clarification. Why are we saying it in months and not 12 months or less? I mean, there is a lot of data.

CO-CHAIR SAKALA: By the time we get the data in --

MEMBER AUSTIN: Yes. If we were just filling out --

CO-CHAIR SAKALA: Eighteen is not enough time.

MEMBER BELL: You need a full year's worth of ICD-10 data.

CO-CHAIR SAKALA: Okay. All right. So, let's vote.

MS. ROBINSON-ECTOR: Voting is open on the proposal. Option 1 is yes; option 2 is no.

(Voting.)

Yes is to agree with Matt's proposal.

Okay. All the votes are in and voting is now closed.

A hundred percent voted yes and zero

voted no. So, that motion passes.

(Applause.)

CO-CHAIR SAKALA: Last comment.

MEMBER WESTHOFF: Well, this is a more generalizing question, being new to the Committee. Has the Committee done anything like this before? And given the larger national need to harmonize and reduce the number of measures, is this new? It seems like a good idea.

And then, second, either having done it in the past or thinking ahead to doing it right now, who is it at NQF who is now tasked with -- you know, there is the burden on the NQF side.

DR. WINKLER: Yes, you're looking at us.

(Laughter.)

We will be staffing this Committee as long as it's in existence and as long as we are. So, this goes on our to-do list to keep everybody reminded. And again, it will become one of the off-cycle activities.

Because we moved to standing committees two to three years ago, we have begun having some of these, "Hey, you know, we couldn't resolve everything today," but we need more data; something is going to change in three months, blah, blah, blah. We want to revisit sometime down the road. So, it is becoming something we see.

The issue around related and competing is huge across all measures, across all topic areas, in terms of really understanding burden and the use of measures. So, this is not unique.

CO-CHAIR SAKALA: Okay. Thank you.

So now, we move on to 0471, which is PC-02 in The Joint Commission's core set, Cesarean -- it says "Section" here -- but it is "Birth" now.

And no one is recused, I believe. Jennifer Moore and I are discussants. And we will begin with an introduction from the developers.

MS. MILTON: Thank you. This is

Celeste again from The Joint Commission.

PC-02, Cesarean Birth, the denominator population is comprised of patients that are having their first live birth and it is a singleton in vertex position, and they have reached term, which means at least 37 weeks f completed gestation or more. Of those patients, then, in the numerator would be those that ended up having a Cesarean delivery.

The goal of this measure is to reduce the number, but not to zero. There's always going to be a Cesarean rate. This is a variation of a primary Cesarean birthrate that hospitals have looked at for years.

So, this is just focusing-in on an area where ACOG especially has made a recommendation that we focus our efforts to make sure that we take a look at variation in practice. The goal is to get this rate close to the Healthy People 2020 of 23.9 percent. We are still above that nationally.

CO-CHAIR SAKALA: Thank you.

MEMBER MOORE: There was only one comment that came out of our Workgroup meeting they think is worth mentioning. During the discussion, we questioned whether it should be classified as an intermediary measure instead of an outcome measure, but that really was the only thing that came out of our discussion to pass along.

DR. WINKLER: Again, I think there is a certain level of philosophy around here. I have thought about this one for years. I can even argue it is a process measure.

So, fundamentally, we let the developer give its assignment. It doesn't have a great deal of impact. There is a huge interest in outcome measures, priority over process measures. So, I do think it qualifies in that kind of dichotomy, but I am not sure it is hugely different, whether it is intermediate outcome and the final outcome is healthy mom and healthy baby or it is a pure outcome itself.

MEMBER MOORE: Okay. Thank you.

CO-CHAIR SAKALA: So, on the evidence, the documentation says no new evidence. We would say that there is continuing affirmation of the same conclusions that were made, that eliminating avoidable, safely avoidable, Cesareans has important benefits for moms and babies, and there are supported ways of doing so. So, if there is no objection, I would take the position that we don't need to re-vote on that one.

And then, on the performance gap, as Celeste mentioned, the Healthy People target is 23.9 percent, and the 2014 data with 1388 hospitals reporting is 26.8 percent. But the variation for this measure is really quite large. So, the performance was 14th percent at the 90th percentile -- wait -- at the 10th percentile, 14 percent, and 40 percent at the 90th percentile. So, there is quite a bit of practice variation out there and, also, just disparities for various kinds of socioeconomic variables.

So, I think we need to vote again on the performance gap, just to see where changes

are and have been. And I would say that we, after a steady rise, we have been plateauing. So, that is kind of the change right now, is stopping that rise. We can talk later that the environment is such that we should be expecting to see this actually turn around. But, right now, I think that is a good beginning, is to stop the rise.

So, any other comments on performance gap?

(No response.)

Okay. So, let's vote on that.

MS. ROBINSON-ECTOR: Voting is now open for performance gap of Measure 0471. Option 1 is high; 2 is moderate; 3 is low, and 4 is insufficient.

(Voting.)

All the votes are in.

Eighty-eight percent voted high; 12 percent voted moderate; zero voted low, and zero voted insufficient. So, for performance gap of Measure 0471, the measure passes.

CO-CHAIR SAKALA: Great.

So, on reliability, I think we do need to take a quick vote. This measure was aligned with the ACOG-led reVITALize Project. So, thus, the new name, for example.

And also, they report an improved ability to identify cases. So, there is a little bit of change there.

Can we pass on this if it is even better than before?

(Laughter.)

DR. WINKLER: Please go ahead and vote.

CO-CHAIR SAKALA: Yes. Okay.

Yes, Jaleel?

MEMBER MAMBARAMBATH: So, I have a comment, a question about the denominator exclusions. One of the exclusions is enrolled in a clinical trial. Now I'm talking in terms of the bigger hospitals. If I can look at my own institution, there are, at least with neonates, in my own hospital there are, at any given time,

there are about 15 to 18 clinical trials going on. So, almost all of the babies will be included in clinical trials. So is the case with many of the NICHD Neonatal Research Network Centers. And I would assume that the MFM Network would also have similar numbers. So, are pulling out a lot of these moms away from the denominator.

CO-CHAIR SAKALA: Celeste, did you have a comment on that?

MS. MILTON: We're actually removing clinical trials in our next version of the manual. So, we were just reporting on the data as it had been collected for this submission. Because we found that the numbers weren't really that large as you looked across the board, when we looked at a 12-month exclusion report. And you have to look at why are they in a clinical trial. Is it really directly related to what we are measuring? So, that was part of it, that we weren't really seeing that there was a lot of variation as a result.

MEMBER MAMBARAMBATH: So, you leave out --

MS. MILTON: Clinical trial, yes.

DR. WINKLER: Okay. I need to ask, Celeste, when does that go into effect?

MS. MILTON: July 1st, 2016 discharges.

DR. WINKLER: But is that reflected in the specs that you submitted to us?

MS. MILTON: For the previous specifications, because they had to be in before we finalized the specifications.

DR. WINKLER: Right. So, if the Committee's okay with that, I am going to ask her to make that update on these specs for where we are now.

MS. MILTON: We can do that.

DR. WINKLER: Thanks.

CO-CHAIR SAKALA: Great. Okay. So, can we have a vote, please, on reliability?

MS. ROBINSON-ECTOR: Voting is now open for reliability of Measure 0471. Option 1

is moderate; 2 is low, and 3 is insufficient.

(Voting.)

All the votes are in and voting is now closed.

Eighty-eight percent voted moderate; 12 percent voted low; zero votes insufficient. So, for reliability of Measure 0471, the measure passes.

CO-CHAIR SAKALA: Thank you.

So, on validity, the developer reports previous and continuing face validity for measure users as well as a website that picks up questions and issues from the field and tries to deal with them in a continuous process of clarification and refinement.

I feel that the exclusions seem appropriate, but, as was just discussed, many of them, so few cases are eliminated, that this seemed to me to be an area where we could look at reducing the burden of collection by really taking out the ones that aren't materially impacting the results.

It was switched to ICD-10 codes using a careful process with checks and verification.

And then, we got some comments, pretty extensive, in the pre-meeting period regarding adjustment for various demographic variables. I don't know if we can bring those data up, but we were provided with some PowerPoints from Elliott with data from 231 California hospitals showing that hospitals with a higher concentration of older moms -- and that is over 35 years -- and higher concentration of moms who had a BMI 30 or higher just before being pregnant were distributed across higher, medium, and lower range NTSV hospitals.

So, what that is suggesting is that there is not a pure risk among women, but it depends on clinical practice. Just to give an example, a woman comes with a BMI of 33. Do you say this woman is at elevated risk if she has a Cesarean and I'm going to work really hard to help her not have one or do you say this woman is headed that way and I'm going to lower my

threshold for going there?

So, that is an example. I believe you are, on July 1st, also eliminating the age bans, is that correct?

MS. MILTON: Yes. That will be effective, again, with July 1st discharges of this year.

CO-CHAIR SAKALA: Okay. And any other comments on that?

DR. MAIN: This gets to really the fundamental issue of what is driving the variation among the hospitals, which is different than among patients. So, when you look at hospitals, what this graph shows, the green dots are hospitals that have low NTSV C-section rates and the red dots are hospitals that have high rates of NTSV C-sections, 35 percent or more.

And you can see on the x-axis is the proportion with high BMIs, and the y-axis is the proportion with high rates of maternal age. There is a general trend that the older the population of nullips, first births, the thinner

the population, and vice versa; you are going to be heavier in your first birth. So that they do have a tendency toward balancing each other out.

But anywhere along the line here, for every red dot, you see green dots next to it. So, a hospital that has the same distribution of high BMIs or high maternal age can have very different rates, which strongly indicates that this is related to practice patterns. It is likely some effect of those are, but you certainly don't want to bake-in the population rates of high BMIs or high maternal age if you have hospitals that can be very well at one end or the other.

And that is really the struggle here, is to tease out what is related to the practice pattern and what is related to the patient. And it is really not the individual patient we are talking about here, and I think that is a super-important point. It is really the practice pattern of the hospital taking care of all the patients that come to that hospital.

CO-CHAIR SAKALA: And I would just like to say that I think baking-in practice patterns where there is significant room for improvement is a problem that we should try to avoid.

Cindy?

MEMBER PELLEGRINI: Can the OBs in the room enlighten me? Are there strong, clear consensus practice guidelines for how to deal with these two populations? In other words, are some of these hospitals adhering to guidelines, whether their rates are high or low, and others are not?

DR. MAIN: No.

MEMBER GOYERT: No. If the question is, are there guidelines for particular labor management, no.

CO-CHAIR SAKALA: Microphone on.

MEMBER JOLLES: All right. Okay. Sorry.

I mean, well, if we broaden your question, is it a guideline specifically on the

elderly nullip or is it a guideline on the obese patient? But the standard guideline that is the root of the CMQCC work and the SMFM, yes, it is all about stopping the overdiagnosis of labor dystocia, and that is what this is all about, are the elderly nullips and the BMI. Okay. Correct? So, yes.

MEMBER PELLEGRINI: We prefer a word other than "elderly". Thank you.

(Laughter.)

CO-CHAIR SAKALA: There are guidelines in terms of obviously -- oh, Sindhu?

MEMBER SRINIVAS: Also to answer that, I mean, I agree there are things that are published as risk factors for having a C-section and how individual clinicians utilize that information to make decisions.

I think Carol pointed out the sort of two versions of how somebody could take that into account, but there's not like guidelines that say, you know, manage labor this way for this person. And I think people just use that

information differently.

I had a different comment, which is -- I don't know if it is appropriate to say here, but just more food for thought -- you know, we talked about a measure yesterday that was about adverse term-birth outcomes. When we talk about lowering the C-section rate, there is a countermeasure to lowering the C-section rate, which is not lowering it too much or inappropriately in certain places where you need to increase adverse neonatal outcomes. You know, lots of measures have a countermeasure that could have a negative consequence. I think the elective delivery one is another one.

And I don't know if there is a precedent or if The Joint Commission is considering sort of kind of pulling together the low-risk C-section rate with a countermeasure that sort of is ensuring that, while we are trying to lower the C-section rate, we are not actually leading to unintended adverse consequences.

DR. MAIN: We are certainly doing that in California, in Oregon and Washington, where these are all in play. I can't speak to where The Joint Commission is going on this.

MS. MILTON: There aren't any plans at this time, but, certainly, this is something we could discuss with our Technical Advisory Panel.

CO-CHAIR SAKALA: Tracy?

MEMBER FLANAGAN: So, I want to echo what Cindy just said about going too low can be just as bad as being too high. I have seen this scatter graph a couple of times.

We in Kaiser Permanente ran our 66,000 births in a one-year or a two-year time period. What we found is multipliers of risk for five conditions. One was age, BMI, race, hypertension, and diabetes. And the biggest factor was pre-pregnancy diabetes increased the risk by 130 percent.

Within our system, while we have a very low rate as a system, we have hospitals that are higher, and we have a very consistent care

pattern because we have 24/7 staffing. I mean, we really have fairly consistent care.

So, while I think Elliott is making the point that there is variation in clinical practice, there may be some real differences based on medical issues that really impact whether babies fit or whether babies tolerate labor, and whether labor goes smoothly in a timely fashion that doesn't exhaust the baby or the placental reserve.

So, while I know that the age adjustment is going away, I actually think that that is okay. What troubles me is what the number is. When purchasers start talking, as has happened in California, that if you are below 23.9, we are not going to pay you, that troubles me a lot.

And so, because that is happening, I really am troubled about not risk-adjusting this. What we are planning on doing is actually taking our own analysis and submitting it to The Joint Commission. We haven't run it by hospitals yet.

We do know we have these factors of risk, and we will see.

CO-CHAIR SAKALA: Thank you.

Nancy?

MEMBER LOWE: Yes, well, I think I am back to the question that was asked about guidelines. I think the literature is very clear -- in fact, I know it is -- that the diagnosis of dystocia during labor is one of the most imprecise, undefined diagnoses that there may be in the whole realm of medical practice.

You know, you can't run a laboratory test. It is in the eye of the beholder what is dystocia during labor. And that is extraordinarily clear in the literature, and I think it is what some of us are trying to help give clinicians better tools to help them with the issue of what is really delayed-labor progress; what is failure to progress; what is failure to wait, all those kinds of issues that really get into this measure of variation by institution.

And so, I just caution us to be very careful about making decisions or assumptions based upon statistical risk, when statistical risk can vary greatly in any one study. And I think Elliott made the point very, very well that there's no bottom-line risk that being older gives you by itself that is reliable and that statistically works across the board.

And so, to me, when we start to tie these kinds of things to outcomes -- and I'm not being very clear what I want to say -- but it is scary when people want to say, "I'm not going to pay you." Because, to me, for this population, 23.8 percent rate is way too high, way too high, for healthy nulliparous women at term, one baby, head down. That is one in every five women, more than one, almost one in every four ends up in the OR. What is wrong with that picture? I think it is a public health problem, that we are doing that much surgery on women to have a baby, which is a physiologic process.

So, enough. But I just want to

caution us about this, not to get so far down the road of tearing things apart statistically that the measure becomes unmeaningful to the public or unmeaningful in terms of really monitoring what is going on nationally in terms of safety and quality.

CO-CHAIR SAKALA: Thank you.

Juliet?

MEMBER NEVINS: So, I am payer, and I certainly would not at anytime support a situation where we paid for a certain percentage of C-section rate. I am also a laborist. You know, last weekend I did a C-section on a baby that was five pounds.

We do have guidelines with respect to, or new guidelines I should say, with respect to what is the definition of arrest of dilation or descent. So, that is available, and there is significant uptake with respect to that.

Most of the obstetricians in obstetric programs are moving towards some sort of a hospitalist model where you do have in-house

obstetricians available. You know, the day of the private OB doctor is sort of going to the wayside. Certainly, in certain sections, regional sections, of the country, that still exists, but certainly in the larger urban centers I would say that most of the hospitals do have 24-hour care.

So, I guess the point, I think that these studies should be risk-adjusted because I do understand the graph, but I also have to agree that it is an art; it can be very subjective. But I would say that there are inherent risks to ending up in the operating room if you have one of these five risk factors, one of them certainly being your BMI. So, that is my slant on it.

And anecdotally -- and I don't know if the other OBs in the room would agree with me -- but no one wants to take a heavy patient to the OR. We want that lady to deliver vaginally, just letting you know. It is much easier for the baby to come out through the vagina.

CO-CHAIR SAKALA: Jaleel?

MEMBER MAMBARAMBATH: I want to caution those comments about balancing risk, because this maybe not relevant to this particular case, but there is more food for thought over here for NQF.

Whenever there is a QI project which comes up for me as a medical director of an EQ, I ask for what are the process measures, what are the outcome measures, and, okay, do you have balancing measures, too.

When we have measures coming up over here, one of the major concerns that we have in the committee is that, hey, is this going to worsen some other outcome? But we don't have a balancing measure included into many of these measures that we have. Should we be including, asking the developers to include a balancing measure in there, along with this as a package?

CO-CHAIR SAKALA: So, and Elliott is going to comment.

DR. MAIN: We are doing major quality improvement efforts on this subject currently in

California. And the balancing measure for this particular measure presented yesterday and that was approved or re-endorsed, which is the healthy term newborn. This is all term patients and the question is how you manage the labor. This is one outcome of the labor management. It is a C-section, of course.

And as I said yesterday, the most outcome is of the baby from birth and that is what is being used around the country and in the western three states and, as we heard yesterday, in other parts of the country as well, as is appropriate for any obstetric intervention.

MEMBER MAMBARAMBATH: Yes, my question is more broad not only for this particular measure but as any measure which comes through to NQF, should we ask them to have a balancing measure associated along with that and come up as a package to NQF, when they come up with the measure.

CO-CHAIR SAKALA: So, I think the staff can take that under advisement.

So, Dianna and Amy and then let's vote on validity after that.

MEMBER JOLLES: I just wanted to speak to another side of the important points brought up. I think sometimes we actually have to question our assumptions and our reality to really move the hockey puck where it is headed. And I will just that while we await Strong Start data, I can speak to the management of a perinatal data registry that I am involved with, where a nulliparous term vertex singleton cesarean rate is race is not a predictor in cesarean. So, the assumption that race has to be where it is, I understand nationwide because of healthcare delivery system issues. We don't have to accept that assumption.

As for preexistent diabetes, PRAMS data and birth certificate data would suggest that population-wide, it is no more than three percent.

So, if you have a system where you are caring for like 20 percent of your patients are

preexistent diabetics, then this is, again, a signal-to-noise issue, where we have got to accept that and move on. But in general, with the amount of unwarranted variation going on, I personally believe that we have to remember that NTSV is risk-adjusted. We are dealing with healthy childbearing women, in general, with a low chronic disease burden, at this point in their lifetime, and that we are fortunate in this beginning of life care group to be able to have this level of equal playing field. Whereas, when you get into the end of life care measures and Medicare, it is a whole different thing where risk adjustment becomes important. But be careful about assuming we need to adjust for certain things where the model of care is predictive of outcome and, in this particular data set I am discussing, hospital and parity were the only independent predictors of cesarean section.

CO-CHAIR SAKALA: Thank you. Amy.

MEMBER BELL: Just a question about

when looking at further development of the measure, if there is any consideration for making a clause in there regarding vaginal -- or C-section delivery is the preferred method of delivery or vaginal delivery is contraindicated.

So, if there is a contraindication for a vaginal delivery, if those patients can be excluded from that measure moving forward.

DR. MAIN: I have looked extensively at other diagnoses what would be contraindicated for vaginal delivery. There is a couple of comments to be made.

One is that they are very rare in a nulliparous population. Conditions such as placenta previa, for example, nulliparous at term. There is a very small number of previous that we meet that. I have also looked at HIV.

And the trouble with both HIV and placenta previa is the coding. There is only 56 cases in all of California that were coded as HIV with the several codes for HIV in pregnancy nulliparous to term, suggesting we were under

coded. But there is not a huge number that would change anybody's rate.

Placenta previa, for example, we had a couple of hospitals that two or three percent of their patients had placenta previa. Half of those were delivered vaginally. Another quarter of those were induced. And the coding was indicative of a placenta previa being present on ultrasound in the first or second trimester that got coded on the delivery chart.

And so as you get into other diagnoses, you get into coding issues that may or may not be real, which is one of the reasons we wanted to keep this as simple as we could with the best quality codes or the best simple indicators. This can also be done using birth certificate data. So, it is used by states and nationally. The NCQS or the national Center for Health Statistics runs this for every state every year and that correlates very, very well. Part of it is that it is clear-cut and simple and it is I think a value to stay that way. Even

though, yes, there is 20-odd percent good reasons to do C-sections, of which some of those fall into place.

Patient choices often raised and that is the area of some percent but it is kind of interesting. It varies greatly by provider and how you talk to your patient and how you discuss the pros and cons. And again, that isn't one that really drives the rate. It is how you treat everybody else in your practice that is desirous of a vaginal delivery that really drives the rate.

CO-CHAIR SAKALA: Thank you. So, let's, in the interest of time, now vote, please.

MEMBER NEVINS: Can I just -- I'm sorry. One quick comment. I will be very, very quick.

So, I just wanted to stress that I certainly understand the dilemma and I certainly appreciate the comment that you made with respect to strident words more and more vaginal deliveries.

And I will just give you an example. I worked as a private practitioner in a very middle class healthy community for seven years. Our C-section rate was like 40 percent because you had it was very subjective. We were limited to the desires of the patients and doing inductions when we weren't supposed to because Mom wanted her baby delivered on a particular day. And so in that situation, you can see where leaving the data pure would tease out C-sections that are being done as the fault of the provider.

However, I know work in a very different environment, where we have 24-hour staff but we have a very sick population. We don't have healthy moms. If you have got ten patients on the labor floor, six of them have preeclampsia.

So, I mean I think that to be fair to the hospitals that have high-risk populations, in my view, I think that it should be risk-adjusted and that is why I am stressing that because it is regional and it depends on the cohort of the

patients that you are dealing with. So, I just wanted to add that.

CO-CHAIR SAKALA: Thank you. I think we need to vote and move on. I'm sorry.

Could we please open up the voting for validity for this Cesarean Birth Measure?

MS. ROBINSON-ECTOR: Voting is now open for validity of Measure 0471. Option 1 is high, 2 is moderate, 3 is low, and 4 is insufficient.

All the votes are in and voting is now closed.

Eighty-eight percent voted high, fifty-four percent voted moderate, eight percent voted low, and zero voted insufficient.

So, for validity of Measure 0471, the measure passes.

CO-CHAIR SAKALA: Thank you. Next is feasibility. And in this case, it is manually extracted from health records by approved vendors among now all hospitals with 300 or more births per year.

And a notable point is that they are working on an eMeasure that will be tested this year. That will be a good addition. And they provide sampling guidelines to their hospitals and have many years of implementation by an increasing number of facilities.

So, are there any other comments on feasibility, including your knowledge from the field?

Okay, I think we have two cards up that are not intended to be comments. So, thank you.

If not, we can vote, please.

MS. ROBINSON-ECTOR: Voting is now open for feasibility of Measure 0471. Option 1 is high, 2 is moderate, 3 is low, and 4 is insufficient.

All the votes are in and voting is now closed.

Fifty-eight percent voted high, forty-two percent voted moderated, zero voted low, and zero voted insufficient.

So, for feasibility of Measure 0471, the measure passes.

CO-CHAIR SAKALA: Thank you. So, moving on to usability and use.

For public reporting, I just would like to ask -- in my view there is a lot of confusion out there and we just saw it with the Consumer Reports release is all this discussion of the cesarean rate. And I feel as if we need to teach all the stakeholders what we are talking about and kind of coalesce around the most meaningful measure among total primary and NTSV. And so would be very eager to have public reporting beyond voluntary circumstances and wondering what the plans are for this, when we could expect it. It was noted in the specs.

DR. OWENS: We are currently trying to figure out how to do it, to be perfectly honest with you. Our public reporting system is set up basically for process measures. We are trying to figure out how to accurately report it publicly, this measure, as well as some others, so that

they make sense to the public. We are working on it. We hear you. We know that it is necessary.

CO-CHAIR SAKALA: Thank you.

And as far as improvement, this measure is used in programs internal to the Joint Commission. It is a population measure. It is included in the Medicaid Child Core Set.

I feel that the ACOG-SMFM recommendations in early 2014 were a really important signal to the field and since then, we are seeing a lot of important work around the AIM bundle and the toolkit that was released last week and that these are around primary but the toolkit says the best measure is the NTSV measure. So, it is really -- and the new or alternate payment models that are coming up and are starting to be used, especially in Medicaid programs for our purposes and the requirement is to be collected in now over 80 percent of hospitals.

So, this measure is really out there and my view is that this issue is the heritor to

elective delivery, in terms of the primary QI focus for our field right now. So, I think this is a really big one around use and usability.

DR. MAIN: My only cautionary note is we certainly don't want to have this be driven down to a very, very low number. I would certainly be supportive of, unfortunately, early elective deliveries being driven to zero and there probably should be three to four to five percent. We don't want to see that with this measure.

But you know there are hospitals in California that are 35, 40, 45, 50, 60, 65 percent on this rate. That is the target.

CO-CHAIR SAKALA: And the other question here is unintended consequences. And again, my view for the next group with the Joint Commission is that unexpected newborn complications would be a great addition to that measure set and would, I think, give a lot of people good peace of mind around this question because we don't know where the right rate is.

And people need to improve in a safe way. Chances are they can really move the needle over time but they need to do it safely. And so we need to have ways to monitor that.

Tracy.

MEMBER FLANAGAN: There was a comment from the presenters about the challenges of public reporting of outcome measures. Could you elaborate on that? Maybe other people at the table don't know what those challenges are.

DR. OWENS: What the issue has to do with is determining the expected rate versus the actual rate and reporting that in such a way that it makes sense to people. That is basically it.

MEMBER FLANAGAN: But why is it more challenging than a process measure?

DR. OWENS: Because we have the system set up without that piece of it and it is a question of adjusting our system.

MEMBER FLANAGAN: Okay.

CO-CHAIR SAKALA: Is that Sindhu? Yes.

MEMBER SRINIVAS: This is just a general question for the Joint Commission. Is there consideration that you would give to even changing the title of the measure just because it is called Cesarean Birth and it makes it seems like it an all cesarean number versus like a low-risk cesarean or whatever in the NTSV or whatever we want.

MS. MILTON: The short name is Cesarean Birth but the full measure description or the full name of the measure goes into that greater detail. And it only has to do with space when you are putting it out there. But if you look at the measure form, it is clearly articulated right below the name of the measure.

CO-CHAIR SAKALA: Any other comments about usability and use? Okay, let's open up for voting, please.

MS. ROBINSON-ECTOR: Voting for usability and use is now open for Measure 0471. Option 1 is high, 2 is moderate, 3 is low, and 4 is insufficient.

All the votes are in and voting is now closed.

Seventy-eight percent voted high, twenty-two percent voted moderate, zero voted low, and zero voted insufficient.

So, for usability and use of Measure 0471, the measure passes.

CO-CHAIR SAKALA: Okay. So, any crucial comments before we turn to our overall question of whether we vote to continue endorsement of this message -- to recommend continued endorsement?

Okay, so voting is now open, please for whether we wish to recommend re-endorsement of this measure.

MS. ROBINSON-ECTOR: Voting is now open for 0471 for recommendation for overall suitability for endorsement. Option 2 is yes -- Option 1 is yes, option 2 is no.

All the votes are in and voting is now closed.

Ninety-six percent voted yes and four

percent voted no.

So, for recommendation for continued endorsement of Measure 0471, the measure passes.

CO-CHAIR SAKALA: Okay, now we are scheduled to take a break at 10:15 and we are scheduled by 10:15 to have done one more measure. So, do people have a sense of what they want to do right now? It is the other measure on cesarean rate.

CO-CHAIR GREGORY: Does anyone want to take a break? How many want to keep going?

CO-CHAIR SAKALA: Okay, can we start by 10:30, please?

(Whereupon, the above-entitled matter went off the record at 10:18 a.m. and resumed at 10:29 a.m.)

CO-CHAIR SAKALA: Let's reconvene, please.

CO-CHAIR GREGORY: Okay, gang, let's reconvene.

CO-CHAIR SAKALA: Okay, so now we are moving on to a new submission, which is number

2892, Birth Risk Cesarean Birth Measure. And no one is recused. And the discussants are Jennifer Moore and Nancy Lowe, and Tracy Flanagan.

And first of all, we will hear from the developer, Dr. San Roman.

DR. SAN ROMAN: Good morning. My name is Gustavo San Roman and I would like to thank you for taking the time to review my measure.

If I could ask one question before I start, I am little confused as to Measure 0471. Did the committee endorse the measure as submitted with the direct standardization age adjustment or was the dropping of the age adjustment what was endorsed?

DR. WINKLER: They dropped the age endorsement for the new version.

DR. SAN ROMAN: Okay, thank you.

Good morning. The Birthrisk Cesarean Birth Measure was developed, in part, as a result of not being able to convince anyone six years ago that the flow in the direct standardization risk adjustment of Measure 0471 would become

problematic.

With that in mind, I set out to develop a better measure and found that research has shown that the physical characteristics of the mother and the size of her baby significantly affect a woman's risk that labor will result in a cesarean birth. Therefore, we need more risk adjustment and not less, if the goal is to assess the effect of the obstetrical care provider. Or in other words, every woman enters into labor with an inherent risk that her labor will end in a cesarean birth based on her physical characteristics and the size of her baby. The goal of the cesarean birth measure is to measure the effect that the obstetrical care provider has on this inherent risk.

Inherent risk is not a new concept but it has had different terminology. In 2003, Dr. Bailit referred to this risk as probability of cesarean delivery and she referred to a hospital's average expected probability as the expected or risk-adjusted cesarean delivery rate.

I bring up Dr. Bailit's work on logistic regression modeling only because the method used in my measure is similar to her work from 2003. Unfortunately, it seems that I do not have her ability to describe my work clearly enough for committee members to see the similarity of my measure through logistic regression modeling, as became apparent by the comments during the last work group phone call.

I believe that a quick comparison would be helpful. In logistic regression modeling, an equation is used to predict the number of expected cesarean deliveries based on the risk factors contained in a population and compares it with the populations actual cesarean delivery rate. In logistic regression modeling, this prediction is created by taking a fixed data set and obtaining an equation from the data set that reflects the effect of each risk factor on prior outcomes.

The equation has a coefficient for each risk factor, which assigns a weight to each

of the risk factors. The equation is then applied to each birth record in the data set in order to calculate the expected cesarean birth rate for that woman. A provider's or hospital's average expected rate for their population of patients determines the expected rate for that provider or hospital.

Once the expected rate is calculated, then a simple comparison of the actual rate the expected rate creates the cesarean birth measure.

The only difference in my measure is that instead of using an equation to calculate the expected rate, my measure uses the cesarean birth rate of a cohort of 100 similar patients to assign the expected rate. The cohort method uses the same mathematical concept that is used to create the equation in logistic regression modeling.

For example, the equation assigns the expected rate based on the weight of each risk factor and the weight of each risk factor is dependent on the actual prior outcomes within the

data set. The cohort method bypasses the equation by assigning the expected rate based on the actual prior outcomes.

The reason that I use the cohort comparison method over logistic regression modeling is that the cohort method is more accurate due to the three limitations of using an equation. First, is that the equation cannot obtain accurate calculations if any of the risk factors do not exhibit linear progression throughout their range. Both fetal weight and maternal weight gain do not exhibit linear progression. Second, is that the equation cannot obtain accurate results at the extremes. And third, is that an equation based on a fixed data set cannot account for changing practice patterns.

Using a cohort of similar patients means that nulliparous patients are only compared to a cohort of nulliparous patients. And multiparous patients are only compared to multiparous patients. And this is the reason why

nulliparous and multiparous patients can be included in the same measure. In fact, the cohort used to assign the expected rate will be similar in parity, onset of labor, fetal weight, maternal pre-pregnancy BMI and maternal age, maternal height, gestational age, and pregnancy weight gain. Being able to provide risk adjustment for eight risk factors is a significant improvement over other measures that only adjust for age or perhaps for nothing at all.

In summary, measuring the effect of the obstetrical care provider is extremely complicated and only a complicated cesarean birth measure will provide accurate results.

As I mentioned in the last phone call, there is a 33-minute PowerPoint presentation on my website that provides additional information and explanation of my measure.

Lastly, also in the last phone call, there was a question about statistical analysis of reliability and validity. The statistical

analysis was done by Dr. E.K. Ahn, who trained at Harvard, Columbia, and Stanford. I gave her the data and the questions that needed to be addressed and she provided the answers. And I was asked to provide her analysis and I have brought copies with me here today.

Thank you.

CO-CHAIR SAKALA: Thank you very much.

So, we will turn to our discussants to begin a discussion of the evidence.

MEMBER LOWE: Just a couple of more orienting factors and thank you, Dr. Roman, on the measure so everyone is clear. The measure is described as being a measure of the effect that the obstetrical care providers' labor management strategies have on their laboring patients' risk for cesarean birth. The target population is limited to women who attempt labor with a singleton vertex pregnancy without a history of prior cesarean birth and give birth between 37 and 42 weeks gestation. And an important point to remember about this measure is it does include

both nulliparous and multiparous women in the measure.

The level of analysis, according to the report, is that the individual clinician or the facility level and the data source is birth certificate data. The numerator is the number of women undergoing cesarean birth and the denominator as described is all women without a history of prior cesarean who attempted labor and gave birth to a single baby in a vertex presentation between 36 weeks, 4 days, and 42 weeks, 3 days.

And as described, the risk adjustment is by cohort comparison to previously recorded births to determine the expected cesarean rate for the target population.

As a new measure, I expected to see evidence of how this particular measure was related to other identified outcomes and that was not provided. The developer's summary of the evidence is that there are many different labor management strategies that have been used over

the years to assist women who are in labor. Some of these strategies can decrease a woman's inherent risk that labor will result in a cesarean birth and others can decrease her inherent risk. And examples were provided, including doing an operative vaginal delivery or an inpatient obstetrical provider as another example.

While I agree that these things that the provider and nurses do -- there was an interesting about 20 years ago that nurses have individual cesarean delivery rates, by the way, primarily ignored but a very interesting analysis -- to increase or decrease a woman's risk for cesarean delivery, it is unclear how this particular calculation does that.

Further, the developer did not specify how this measure uniquely captures those dynamic relationships, particularly the dynamic relationships between maternal care truistics and provider decisionmaking because it is not simply additive. It would be multiplicative in some

cases. Nor did it provide evidence that this outcome measure has any empirical relationships to various specific processes of care.

So, my personal evaluation is that it is a no-pass on the evidence.

CO-CHAIR SAKALA: So, I think that actually what you are discussing is a little later down the road in our discussion. I have heard you say it should be the same as the other evidence around --

DR. WINKLER: Well, it is just I would say in terms of we just talked about a measure that was about cesarean section rates. And this is a measure about cesarean section rates. They are very different. But nonetheless, it is still measuring the same concept. And for evidence, we are talking about the concept, not the specifics of the measure, per se. That will come later on when you are looking more at the reliability and validity of the specifics of this particular measure.

This is an outcome measure. If you

want to, we can quibble over it. It is a new media or an outcome measure. And so really the requirement for evidence for outcome measure is are there processes, structures processes or other activities care that can influence the outcome and that is really the question for the criteria for an outcome measure.

CO-CHAIR SAKALA: So, hold that thought, Nancy, about other concerns. But let us first assess whether the evidence is there for improvability.

MEMBER LOWE: So, then we are talking at the very high level, not this specific measure, but we are talking at the high level.

CO-CHAIR SAKALA: Yes, about the concept right now.

MEMBER LOWE: Okay.

CO-CHAIR SAKALA: And we will need to vote on this, even though it is the same as what we voted on.

MEMBER LOWE: It doesn't matter. This is a new measure. It gets evaluated against

everything.

CO-CHAIR SAKALA: Okay. So, shall we -- somebody down there go ahead?

MEMBER FLANAGAN: It was my understanding in evaluating evidence that not only was there a theoretical idea that you could lower C-section but there was in fact an intervention and a tested intervention that showed that. And this measure makes the claim that labor management affects C-section but there is no testing of this anywhere.

So, I mean if you feel that my comments are not relevant to the evidence here, I will say it again, but honestly, I think that almost every other measure we have evaluated has some published data that shows something of direct relevance to the measure and I don't see that cited here.

CO-CHAIR SAKALA: No, you don't have to do that. I just wonder if we should wait until we get to the right place for this and just --

DR. WINKLER: It will be helpful if we go through the criteria appropriately but, as I said, evidence is around the topic area, not necessarily the specifics which is why, if you noticed, we had several measures on infection. They weren't the same but the evidence ultimately will be the same for supporting both of them. Similarly, I think we have got two measures that address the issue around cesarean section. So, the evidence should be similar, even though some of the details, once we get into the specifics of the measure, and the differences are where you are going to have some divergence.

CO-CHAIR SAKALA: Diana, do you have a comment about that first section? Okay.

MEMBER JOLLES: I just wanted to mention the, and I apologize for not knowing how to pronounce her name, Kozhimannil article that was published in 2013. I'm concerned about its lack of inclusion in the summary of evidence.

No one here can argue that risk stratification and risk adjustment is important

in this outcome and that there has been incredible epidemiologic large database studies that show that these things affect cesarean rate. However, importantly, what has been shown is that actually low-risk women are more affected by unwarranted variations in care and supply-sensitive variation than women with risk factors. And just because part of what we are supposed to be doing is talking about our personal experience and bringing our content expertise to the table, I was just asked about Tuba City, the section rate and their diabetes rate. Thirty percent diabetes rate, primary section rate of nine percent.

And if you look at the hospitals that achieved the lowest cesarean section rate in the Consumer Reports articles over the years, they have very much held a high-risk population.

So, I am concerned about evidence. I'm not discounting the evidence here. I'm just saying that there is more evidence out there and that the issue is complex.

CO-CHAIR SAKALA: Okay. So, the evidence out there supports the importance of this measure and the opportunity to improve.

Can we take a vote, please, on whether the evidence is there for this measure concept?

MS. ROBINSON-ECTOR: Voting is now open for evidence of Measure 2892.

All the 27 votes are in and voting is now closed.

Ninety-six percent voted yes and four percent voted no.

So, for evidence of Measure 2892, the measure passes.

CO-CHAIR SAKALA: Okay and I think probably our discussants would agree that the opportunity for improvement is there. That is the nature of the conversation that we have had.

DR. WINKLER: Yes, but we do want to look at the specific data generated by this measure because now we are talking about some specifics, if at all possible. Because, in all honesty, it is quite a different measure as you

start adding in the multiparous patients. You have a much different denominator population.

CO-CHAIR SAKALA: Okay. So, do the discussants want to add anything else to what you have said so far?

Okay, Matt.

MEMBER AUSTIN: Yes, I noticed that the data provided I think are from 2005 to 2007. Do you have any more recent data that would reflect what variation is? And also I think the data were just from one state as well.

DR. SAN ROMAN: Correct. That is the only data that I have.

CO-CHAIR SAKALA: And some of the other data were excluding New York City from New York State. Could you talk about your population that you are adjusting to?

DR. SAN ROMAN: Sure. The data was obtained from New York State Department of Health. And there is two different systems, at least at the time that I requested the data between the city hospitals and the rest of the

state. So, the Department of Health had access to the rest of New York State, not including the city hospitals. So, the data that I have is from hospitals outside of the five boroughs.

CO-CHAIR SAKALA: Okay. Matt, do you have another comment? Okay.

Other comments? Jennifer.

MEMBER BAILIT: Hi, so I have lots of issues with this measure but let me start on the one that I think is relevant for performance, which is that you say that the individual has an appropriate level of evaluation for this and we have shown that pretty much -- and I do have, as you have mentioned, a lot of experience with logistic models for this -- that it is inherently unstable to look at the individual because there is never enough numbers with any one person to be able to get confidence intervals around the expected-to-observed rates to get any sort of stability and to know whether are falling outside of.

So, can you tell us why yours would be

different than any other model, in terms of numbers, since it is essentially a math problem with a small problems when you get down to the level of the individual?

So, I guess that is my question and my comment.

DR. SAN ROMAN: That's correct. We always worry about small numbers when are looking at the individual and some doctors perform more births than others. And some of the data that I have brought into the submission show a hospital that has 86 providers in it. And we do the statistical analysis comparing the providers to the average. And if they are statistically significantly different from the average, based on their numbers and their result, then we could say that. If their numbers are small, they would not be statistically significant.

MEMBER BAILIT: Right but the problem is, if they have small numbers, they will never fall outside the statistical significance. So, your confidence intervals are so wide as to be

meaningless, typically.

DR. SAN ROMAN: Right. For those providers, there will be some providers have larger numbers and the confidence intervals will give you the ability to determine that are they significantly different or not.

I think what you can see at the provider level is that there is a hospital that I put in the submission that the hospital falls out of the confidence interval as a hospital that is doing too many cesarean births.

But if you look at the 90 or so providers that are in that hospital, 80 percent of them don't fall out. So, you have got 20 percent of the providers who are doing about 25 percent of the deliveries in that hospital that that is where our problem is in that hospital, not the whole hospital.

MEMBER BAILIT: I would argue that your problem is small numbers for the other 80 percent.

DR. SAN ROMAN: Okay.

CO-CHAIR SAKALA: Kim.

CO-CHAIR GREGORY: I would just agree. I agree with Jennifer.

CO-CHAIR SAKALA: Naomi.

MEMBER SCHAPIRO: So, here is my difficulty about this. It seems like there is definitely a performance gap like globally. But just looking at this measure, I have some concerns about the narrowness in which the data were drawn and especially being drawn so long ago.

For example, we have been talking about the transition from the individual obstetrician coming and delivering the baby to laborists. And that has been really accelerated, I think, in the last maybe ten years. I mean it is not exactly my field but it is really a big trend now and it hasn't been then.

So, if you have laborists and they do come and go. So, you have laborists who are on for a certain amount of time and then maybe somebody else comes on shift in a way and

delivers the baby. That is really the combination of two people's decisions during the labor management. And so if you are just going to put that on the person who actually did the C-section, it might not be fair.

So, I am just having some trouble. And again, I'm not expert. I am really more with the teens trying to keep them from getting pregnant but I am just having some trouble wrapping my head around this as a concept in this particular measure.

CO-CHAIR SAKALA: Yes.

DR. SAN ROMAN: So, let me just add something to that statement because I think that is very important.

What I have been doing is now the National Vital Statistics has made available national data. So, the 2011 birth certificate data is available and I have pulled it into the data set. So, the data set now has 2.4 million deliveries in it.

So, now there is a greater volume of

cohorts that could be used to find the cohort of 100 patients. However, that data is not hospital-specific. It is not provider-specific. It is not even state-specific. So, I can't do any analysis on that data itself, other than providing a much more robust data set to find the 100 patients that we are going to compare to.

I always find in obstetrics, and I am one of those dinosaurs in solo practice. I have been doing it 26 years but I find that it is the person who initiates the labor management plan that really should take the responsibility for how that labor is managed. And that could be the doctor who comes in and admits the patients, says we need to induce you because whatever. The guy who comes on shift now is stuck with whatever his predecessor has given him. So, in my mind, that is really the person who should carry the weight.

And this is a new measure. So, I just have the data that I have but the goal to move forward is not to look at the individual provider because it is a team of providers. I think the

lowest level you can look at is whatever that on-call group is. If there is a group of four doctors that share call, I think that really would be the lowest level that you can look at fairly.

Because even in our own institution, we see doctors that are within a group who always pass off the patients to somebody else and the somebody else ends up doing the C-section.

The other is if you look at a team or an on-call group, you can include midwives in this measure. You can include birthing centers that initiate a labor management plan and then they will actually have a C-section rate because that patient who needed to get transferred to the hospital now would count against where that labor management was initiated.

CO-CHAIR SAKALA: Nancy.

MEMBER LOWE: Yes, I'm struggling with how to move us forward a little bit with this because I think that in all the time, which was considerable, I spent on this measure, when I

think of our purpose in NQF, I am not sure how, at this point, this complex of a measure could have much usability or feasibility for public information that would be interpretable, for want of a better word, by the general public.

And I am also struggling with the fact that the analyses presented, the seven maternal factors and so forth, are based upon historical data from the data set. So, risk is calculated on the basis of what happened to that woman, which is, indeed, a reflection of what the provider did.

So, to me, there is another step in the analysis, which is teasing apart the provider from the woman, which is multi-level modeling is what that really is, from a statistical standpoint.

So, I think there is more work to be done. For example, publishing your work, where it gets peer-reviewed by the scientific community and showing how the measure, indeed, is related to outcome would help us a lot.

So, I mean that is just where I am and I am struggling with that whole piece.

CO-CHAIR SAKALA: Thank you. So, could we have a quick response, please and then Naomi and Juliet. And then let's plan to vote on the question of opportunity for improvement. We just need to go through the criteria to get to the right point, so that we can weigh it.

MEMBER LOWE: Yes, right.

DR. SAN ROMAN: Quick response. Totally agree that my work should be published. I presented it to a dozen journals and it is complicated. And as you see here in this committee, it is not all that easy to grasp and I am not the best person, perhaps, presenting that information or writing it on paper. But the responses I got from editors was it is good work but maybe it is not important enough to be published. But I have attempted to publish the work.

CO-CHAIR SAKALA: Thank you. Juliet, do you have a final comment before we vote on

opportunity for improvement?

MEMBER NEVINS: I had a question about the cohort comparison. I think Nancy just answered that. So, I will just end by saying that I agree with your comments with respect to the complexity of this model and the potential for use.

CO-CHAIR SAKALA: Okay. So, could we open the voting please for opportunity for improvement using this measure?

MS. ROBINSON-ECTOR: Voting is now open for performance gap for Measure 2892.

All the votes are in and voting is now closed.

Seven percent voted high, twenty-six percent voted moderate, twenty-eight percent voted low, and nineteen percent voted insufficient. So, this would be a grey zone.

DR. WINKLER: No, it's not. So, the measure fails. And one thing I want to make clear -- because we will stop at this point -- one thing I just want to understand is the

rationale. And is it because essentially the data we have that is used in this measure is from almost a decade ago and the performance is old, in terms of this particular criteria? I know we have talked about a lot of other things. I am just wanting to be sure I can explain your vote on this criteria.

MEMBER BAILIT: So, I think this is a couple-fold. The question is can we change performance gaps with this measure. My answer is now for a couple of reasons. One is that it is focused on the individual or the practice level and not at a high enough level. Two, the methods are non-standard. They are close to the standard, sort of standard ways that people do this but they are off enough and they are complex enough -- and I will be honest, I stopped publishing the stuff that you referred to because Elliott came out with these NTSV and it was so much cleaner and crisper and more usable that I just stopped.

Yours is even more complex. So, I

think the ability to explain it -- if you can't explain it to this group, your average 19-year-old having a baby is going to have a really hard time with it.

So, to the extent that I think this measure can move us forward to change the problems that we have, I'm concerned.

DR. WINKLER: Any other comments? I just want to be able to reflect the reasoning behind the vote. That's all.

CO-CHAIR SAKALA: Matt, your rationale?

MEMBER AUSTIN: Yes, my rationale was that the data are almost a decade old now and that they were data just for one state. And I think we have seen that practice patterns can vary by region of the country. And so for me, it was a very narrow snapshot of what data could look like.

CO-CHAIR SAKALA: Nancy?

MEMBER LOWE: Yes, I think that the performance gap that is represented is

represented from this very focus from 2005 to 2007 from New York and I think what we are reacting to is our lack of confidence in this measure to move that a performance gap in the cesarean delivery rate. And I am not sure if that is the exact question we are supposed to answer, Reva, if we are a little bit beyond that one question of the performance gap because I think we all agree there is a performance gap in cesarean delivery.

DR. WINKLER: But you're right. This one is, we are talking about this particular tool for understanding that gap. And so you are right, there is an influence of how the tool work --

MEMBER LOWE: Yes.

DR. WINKLER: -- for interpreting what the actual results show us.

Cindy, did you have a comment?

MEMBER PELLEGRINI: Yes, I am hoping you can clarify. And I think you just answered part of this. But I am perplexed and a little

troubled that the vote on this is different from the vote on the previous measure. Where on the previous measure we said yes, big gap, big problem and here we are like well, maybe not so much.

I understand that there is the influence of the tools here but you said on the evidence we are dealing solely conceptually. Here, we are dealing -- I just want to be clear that we are talking about the performance gap no longer just conceptually with regard to all C-section and all that. Now, we are talking about specifically whether this tool is useful for closing that gap or addressing it.

DR. WINKLER: It is both the tool but also the data that was presented by the developer to make the case. And so we are not asking you to go elsewhere and look at it.

Certainly, for a maintenance measure if it were, we would absolutely would want to see data from the use the measure.

MEMBER PELLEGRINI: Right.

DR. WINKLER: And so when we are able to have that on the initial, that is also good because, again, it is the use of that particular tool to do the measurement to collect the data.

MEMBER PELLEGRINI: Right. Like I mean I would have no trouble -- I hear the concerns. I think they are largely on validity, feasibility, usability. I think we look inconsistent by saying the evidence may or may not show a performance gap when we all know there is a performance gap in this area.

CO-CHAIR SAKALA: Cindy, I just want to read the Pathway Guide that we got for this process. The first bullet here is briefly describe any data presented on current performance using this measure. Is there opportunity for improvement?

MEMBER PELLEGRINI: On other measures, we have accepted the other similar evidence or past evidence.

DR. WINKLER: I think Cindy is raising an important point for the committee to consider

in terms of is that the important part. But remember that you are asked to look at the data that was presented in front of you. And if your concern was with that data, that is a legitimate concern as well.

CO-CHAIR SAKALA: So, Naomi?

MEMBER SCHAPIRO: Yes, I mean in a way I was sort of ceding to the expertise of the folks in the room who have really looked at this in-depth. But to me what actually solidified it was that you said you hadn't been able to get it published in any peer review journals. And I feel like if there had been a history of publication and then other people using the tool and some evidence that it actually was really helpful, I would have been more predisposed.

I voted insufficient because I just felt like there wasn't enough evidence that this could really add to the discussion in the measurement of gaps.

CO-CHAIR SAKALA: Okay, and one more. Matt? Oh, Jennifer, too.

MEMBER BAILIT: No, that's okay.

MEMBER AUSTIN: I mean I guess I looked at it as sort of directed, which is when I looked at the Joint Commission measure they presented data from 2011 through 2014. Those data represent national data.

Agreed that where we are at in the healthcare space is no different with this other measure but the data presented there was for 2005 to 2007 and was just one state.

So, for me, based on the instructions we were given, that is how I was evaluating the situation.

CO-CHAIR SAKALA: Yes, so is everyone comfortable with staying where we are right now? One dissent, maybe.

Okay, thank you.

Kim, we are going to turn it over to you now.

CO-CHAIR GREGORY: The folks from Children's Hospital of Philadelphia are tied up and so we need to --

DR. LORCH: So, I am here but I will probably have to leave in about 15 minute at the very latest.

CO-CHAIR GREGORY: Okay.

DR. LORCH: I have got a moderating session to go to.

CO-CHAIR GREGORY: And Scott, when will you be finished?

DR. LORCH: If I push that late, I mean I could probably go to 11:45 and be available for about 20 minutes then. Then, I am tied up until you guys are done as well.

Whatever is easiest for you guys.

CO-CHAIR GREGORY: Okay, go for it. We'll go for it now.

DR. WINKLER: I think we are going to go for it.

CO-CHAIR SAKALA: And there is a request for you to speak up a little louder, please.

DR. LORCH: Okay, I'm on my cell phone, so I apologize for that. I will do the

best that I can. I will be quick on the presentation of the initial measure that many people have that in front of them and then spend a little bit of time answering some of the questions that were brought up in the work group meeting that was held a couple of weeks ago.

This is a measure of neonatal all-cause readmission rates, which parallels other measures previously endorsed by NQF, including a pediatric all-cause readmission rate and several readmission rates in the adult literature.

The goal of this measure was to evaluate potentially different aspects of care quality from the inpatient/outpatient side, including transitions of care and education of high-risk families.

Data from our group and data that we presented in this application suggests there is approximately a 200 to 250 percent gap in readmission rates at 30 days after discharge between hospitals, when you have similar patients of dissimilar gestational ages.

Infants, we propose a risk-adjusted model as much for face validity as anything else. Data suggests that younger gestational age and the presence of a chronic complication of preterm birth, including necrotizing enterocolitis, intracranial hemorrhage and/or bronchopulmonary dysplasia does raise the likelihood of a readmission at the patient level. However, data that we have done does not support well, really supports a marginal change in the risk-adjusted rate of readmission compared to unadjusted rates at the hospital or state level.

However, because most neonatologists would like to see everything risk-adjusted, at least by gestational age, we proposed a risk-adjusted model weighing the challenges of an added complexity of a model of that nature compared to having just an unadjusted measure.

To answer some of the questions that were raised in the work group, the choice of gestational age in this measure was infants born at 23 to 34 weeks. That measure, that time

period was chosen to ensure infants who died by gestational age categories would be almost 100 percent of the unit in the country admitted automatically to that unit.

Infants below 23 weeks, there is high variation in the aggressiveness of resuscitation between units, leading us to be very leery about including such gestational ages into any sort of measure.

And infants beyond 34 weeks, while having a nontrivial and indicated a very elevated readmission rate, run into the problem of the reasons for such readmission.

So, for infants 35 weeks and above, about 80 percent of the readmissions are secondary to jaundice and hyperbilirubinemia, where there is some controversy about what an acceptable readmission rate would be for those infants, compared to needing to keep those infants in the hospital for one, two, or three extra days for further observation.

CO-CHAIR GREGORY: Hello?

MS. ROBINSON-ECTOR: Scott, are you still there?

MS. THEBERGE: Operator, did we lose Scott?

OPERATOR: Yes, his line disconnected.

CO-CHAIR GREGORY: Okay, discussant, you want to start with a conversation about the evidence?

MEMBER YOUNG: So, I will start with a summary for this. This is essentially NICU graduates from the ages of 23 to 34 weeks, who are readmitted within 30 days of initial NICU discharge, all-comers.

And the evidence that they provided for us was about a ten-year set of data from California and it was listed in Appendix 1, in which it showed an incredibly wild variation in readmission rates.

CO-CHAIR GREGORY: This is actually between 23 and 34 weeks.

MEMBER YOUNG: Sorry, between 23 and 34. I apologize. I misspoke. Yes, NICU

graduates between 23 and 34 weeks within 30 days of NICU discharge, their readmission rate.

CO-CHAIR GREGORY: So, from an evidence perspective.

MEMBER YOUNG: So, my initial take on the evidence presented was that this -- looking at neonatal readmission rates or NICU grad readmission rates was varied among the state of California. That is just the one section.

And then there was some additional data presented from New York and Utah, as well, that I actually was able to find in the literature and their rates were even more varied from just California to New York.

So, the evidence out there is that there is a high level of readmission rate but the variance is so wild that it is very difficult to assess whether this measure may or may not provide any remedy.

DR. WINKLER: I just want to remind you in terms of what the questions are around evidence. This is an outcome measure. And,

therefore, what you are looking for is are there any structures process of care that can influence the outcome.

The data on the actual results is around the gap opportunity for improvement.

So, you have got two different things. So, it can get conflated but as we vote on them sequentially, it is good to keep in mind what is what.

CO-CHAIR GREGORY: Sindhu?

MEMBER SRINIVAS: It seems like the evidence for this measure is similar to the evidence for hospital, like adult, as I think Scott was alluding to, the adult readmission rate, and the idea that you improve transitions of care on one hand to offset the time of discharge could potentially alter readmission rates that are unnecessary or avoidable.

CO-CHAIR GREGORY: And he also alluded to the fact that there is already a pediatric measure.

So, is everyone agreeable with being

able to vote on the evidence?

MS. ROBINSON-ECTOR: Voting is now open on Measure 2893 for evidence.

(Voting)

CO-CHAIR GREGORY: And then if we can have a --

MEMBER YOUNG: So for -- I'm sorry, we're doing the next section, right? After evidence?

CO-CHAIR GREGORY: No, we have to -- she has to tell us if it was approved or not.

MEMBER YOUNG: Oh.

MS. ROBINSON-ECTOR: Yes. We are still missing one vote. We need 27. Thank you.

(Voting)

CO-CHAIR GREGORY: Scott, hold on. I think we hear that you're back.

MS. ROBINSON-ECTOR: Okay. If everyone could try to revote one more time, please. That would be great.

(Voting)

MS. ROBINSON-ECTOR: Okay. Great.

Twenty-seven votes are in and voting is now closed. Ninety-six percent voted yes. And four percent voted no.

So for evidence, Measure 2893, the measure passes.

CO-CHAIR GREGORY: So operator, is Scott back?

OPERATOR: He has not joined the phone line.

CO-CHAIR GREGORY: Okay. Then we're going to keep going.

And now we do gap.

MEMBER YOUNG: So, in terms of gap, what the author was saying is that there are -- there is a gap as mentioned previously.

And that gap is mostly due to the quality of discharge planning, antibiotic use during the inpatient stay. And the quality of outpatient care.

And that in terms of disparities, there was a high rate of African-American, black individuals having higher rates of readmission.

DR. LORCH: This is Scott Lorch, I'm back.

CO-CHAIR GREGORY: Hi Scott. We're actually moving through voting. And --

DR. LORCH: Okay. That's fine.

CO-CHAIR GREGORY: Actually, if you could stay on the line, if there are questions that come up. Just so that you know --

DR. LORCH: Sure.

CO-CHAIR GREGORY: The first vote was based on the evidence. And it was a vote in favor that there's potential structural process, or variables that could be -- that could influence the outcome.

And therefore, we're moving through, we are now discussing whether there's a gap. And the discussants have indicated that there is one. And we are -- unless there are any objections, we are ready to vote on that. No? Okay.

MEMBER SHEA: I just have a question. In terms of the data that you presented, if there's any more recent data? You presented data

from 2006 to 2009.

DR. LORCH: So, we are obtaining the data. Yes. This is a difficult data set to get research on.

And so, yes, there will be more data. California is the best data source for this. But, it will take a little bit of time from that standpoint.

National data set specialists like NIS and KID don't have the readmission slag, which allow us to have more recent data more readily available.

But the data do exist for more recent data. We just have to now finalize the obtaining of the data and run the same analysis.

CO-CHAIR GREGORY: The references are as current as 2013 though.

DR. LORCH: Well actually --

CO-CHAIR GREGORY: The data references.

DR. LORCH: I didn't hear that.

CO-CHAIR GREGORY: I made a comment

that the references cited are as current as 2013.

DR. LORCH: That is correct.

CO-CHAIR GREGORY: So, if there are no objections, we are going to vote on whether or not there's a performance gap.

MS. ROBINSON-ECTOR: Voting is now open for performance gap, Measure 2893. Option one is high, two is moderate, three is low, and four is insufficient.

(Voting)

MS. ROBINSON-ECTOR: All the votes are in and voting is now closed. Fifty-four percent voted high. Forty-two percent voted moderate. Four percent voted low. And zero voted insufficient.

So, for performance gap of Measure 2893, the measure passes.

CO-CHAIR GREGORY: So, now we're going to move to a discussion on reliability.

MEMBER SHEA: So in terms of reliability, this is -- it's administrative claims data, electronic clinical data and

electronic health record at the level of the facility and of the State.

And when I was reviewing this measure, I thought about, actually it's more the feasibility of actually collecting this data at the state and facility level. But, looking at, you know, border States and you know, where do you put the child?

Do you put them in the hospital where they delivered? Or put them in the State that they reside in? And that there would be an issue in terms of the State level data.

And I'll leave it at that. And you know, get to feasibility later.

CO-CHAIR GREGORY: Any other comments? Sheila?

MEMBER OWENS-COLLINS: Let's see, I have a comment. I have a comment and a question. I think this is a very relevant measure.

And speaking from the Medicare side, these -- the micro preemies or the incidence in the State of Maryland is increasing. And not

only are they high cost in the nursery, but they're also high cost in the first year of life. We're looking at that.

And I would bet that they also have frequent ER visits. But, I know that's not a progress measure.

But, I was wondering why congenital anomalies was excluded? Because that is a function of subspecialty care.

And I think at some point during the conference call the availability of subspecialty care as well as primary care providers was mentioned as a factor. And the availability of these providers as a factor in reducing the admission rate.

DR. LORCH: Sure. I'm happy to answer that question very briefly.

We felt that the distribution of congenital anomaly patients between hospitals were dramatically different then the distribution of the typical premature infant. And we were uncertain whether risk adjustment would

adequately be able to adjust for that issue.

And so we felt like it was a safer measure to just look at the prematurely born infants without a congenital anomaly. Many of the congenital anomaly patients are born closer to term.

But even if they're not, they're more of a late pre-term period. It's just was a little more of a reliable measure when we took out those patients.

I readily agree with what your comments are. It's just the reliability became a little more challenging by including those patients in the model.

MEMBER OWENS-COLLINS: Okay. Thank you.

CO-CHAIR GREGORY: Any other comments? Yes, Jaleel?

MEMBER MAMBARAMBATH: Yes. Like Karen, I'm still a little -- finding it difficult where to place this. Whether to place this in reliability or feasibility.

But, since -- I thought I'd bring it up over here because this question might spill over to feasibility as well. But, since the level of analysis is at the facility level or the State level, I thought I would bring it up over here.

So, this measure was tested with hospital data from the California patient discharge data, emergency department, ambulatory surgery and a birth cohort linked with Vital Statistics.

Now, -- and the developer and some of the Committee members also brought this up that there are similarities between this measure and measures for readmission for pediatric patients and for adult patients.

But, I think this is slightly different because there are two different kinds of NICUs. One is the delivery hospital NICU where there is a mother and baby therein. And there are also freestanding children's hospital NICUs.

So, whenever a baby gets discharged from a facility which only takes care of mothers and babies, those babies do not get readmitted to the same hospital. They get readmitted to a pediatric hospital.

So, if you are looking at this at a facility level, it might not be feasible. So, I don't know whether this is -- if it's possible to replicate this in other States and nationwide.

DR. LORCH: So we also have the data for other States. Even some more administrative data. When I think if you want to see, if the question is whether the hospitals themselves can obtain the data, they can. Many of them do not.

It doesn't take one of two issues. One is a more regionalized electronic health record if one is going to use such information. Currently most hospitals do follow up with their patients by phone if they're not going to readmit.

Because even at a children's hospital, they may not be readmitted at the same hospital.

That definitely is a situation that does not always happen.

They could definitely be readmitted at a nearby community hospital such -- depending again on the -- where they live. And what the preference of the outpatient provider was to that.

So, it's definite at the State level from the administrative perspective that type of linkage allows you to see -- and both at that level as well as the insurance level, you can see where the patients were readmitted.

If the data are being collected at the level of the hospitals themselves, what most facilities are currently doing is doing primary data collection of their patients after they come home.

CO-CHAIR GREGORY: Raj?

MEMBER WADHAWAN: I had similar concerns as what Jaleel raised. Because it's not just the children's hospital issue. It's also a very regional issue.

And I don't know how you would use this nationally. Because there are communities with three children's hospitals for a million population. And there are communities with one children's hospital and four million population.

So, where patients get admitted can be quite variable. And who do you assign that to? And how do you capture all of those patients, is I think is a real reliability and feasibility issue in this.

CO-CHAIR GREGORY: Janet?

MEMBER YOUNG: I was going to echo that sentiment. I happen to work for a large health system in Southwest Virginia that also includes a catchment area from West Virginia and Tennessee.

We are the regional children's hospital as well. And so several of those are states apart.

So, for us, this measure would be from a reliability factor, would be quite difficult. And also, it's a significant threat to validity

in terms of readmitting to a completely different institution in a different State away.

CO-CHAIR GREGORY: Sheila?

MEMBER OWENS-COLLINS: So, you know, we're in the era of HIEs. And so, I think that's going to help with the information flow between facilities as a first point.

And then the second point, especially for Medicaid, the Medicaid population, you know, I would think that the MCOs and care managers would be better apt and able to track these patients across facilities.

So, that may determine at what level you want to make this metric.

CO-CHAIR GREGORY: So, I would like to ask a question to the developer. Would not these patients be picked up under the pediatric readmission code because they're pediatric cases?

DR. LORCH: That's a great question. And I know the developers rather well of the pediatric one. They were, I don't want to say explicitly excluded, but they were sort of

excluded from the measure.

The challenge is identifying them in the data set explicitly. Which can become a difficult because gestational age is not necessarily in the data that are presented for the pediatric measure itself.

So, we're actually working on that very question of how well it's being captured. But, both of us believe that it is a separate measure itself. I'll leave it at that.

CO-CHAIR GREGORY: Okay. I understand that the developer is going to have to leave in about five minutes. So, I'm going to take an opportunity to allow the panel to ask questions that might be a little bit out of order so that he has the opportunity to respond.

So, Jaleel, did you have any other issues that you wanted to talk about?

MEMBER MAMBARAMBATH: Yes. The question is whether there are structures and processes available right now as we speak to capture this information? I have not got a clear

answer on that from the developer.

DR. LORCH: At what level? And so, there definitely are some at an insurance level. There are at the facility level. They vary between facilities.

As I said, most of them do rely on primary data collection and follow up of the families themself. And/or contacting the primary care physician who's maybe seen the patient after discharge.

Explicit HIE types of capture rate are in process. But there's nothing currently done that I'm aware of that is being rolled out at the present time.

And that obviously does include going with the border State question that arise from that.

MEMBER MAMBARAMBATH: I'm still not convinced about this. I just want to give an example of my own institution.

We have about four NICUs. Three of which are delivery NICUs. And within those three

NICUs we have about 20 thousand deliveries that happen.

And I am really interested in getting this information. And have not had a chance to get this information about readmissions to different children's hospitals. There are three different children's hospitals within the city. And multiple other community hospitals.

So, I am not convince that yes, that information is currently available. And also, from the insurance point of view, some of these are insurance still -- which is still not -- it's still pending. Medicaid is pending and things like that when they're discharged from the hospital

So, I'm not sure whether we'll be able to get that information from the insurance companies. And insurance companies change as has been mentioned in the document by the developer himself.

And the other issues which has been part of that already. I don't want to redirect

that again.

CO-CHAIR GREGORY: And for a point of clarification, you intent would be that this would always be obtained through a linked data set?

DR. LORCH: That is what the information that we currently have for this measure. I'll leave it at that.

But yes, I think from a statewide perspective that is the most efficient way to get the information. And then deliver it back to the hospitals themselves.

CO-CHAIR GREGORY: Greg?

MEMBER GOYERT: Just to clarify. So, this measure as it stands now from the developer's perspective is going to require new data collection, correct?

DR. LORCH: And/or linkage of data that many States do not give, yes.

CO-CHAIR GREGORY: Sheila?

MEMBER OWENS-COLLINS: I think this is extremely important. Because this is a very

costly issue. Especially for health plans.

And health plans are, you know, at least the ones that I have been working at now and have worked at, are very interested in this data. And they, you know, they are looking at it.

You know, and it is possible that there could be some cooperation at the State level to look at this issue. Because, I mean, these babies are just extremely costly. Not only in the NICU, but after they go home.

And so, I think it is important that we try to figure out which individuals will work to keep them at home and out of the ER as well as out of the hospital.

CO-CHAIR GREGORY: Okay. I'm going to take three, looks like four more questions and comments. And then I'm going to pull the agenda back to order.

Jennifer?

MEMBER MOORE: So, I actually have a question, a clarifying question to help me better

understand this measure.

So, when I was at AHRQ, I worked on a project with Anne Elixhauser and Claudia Steiner with the HCUP data. We produced Staff Brief Number 153 in 2013 on readmissions to U.S. hospitals by diagnosis.

And we looked at discharge data and billing data linking pediatric patients to these readmissions. So, I'm struggling to understand how this measure is different. And I don't know.

And I apologize. I don't know if AHRQ has, or maybe Reva knows, an NQF endorsed measure.

DR. WINKLER: No, they don't.

MEMBER MOORE: Okay. So, it's just based on the HCUP analysis we do. So, I guess I'm struggling, what is the piece of information that we aren't getting already?

And I'm directing it to you too, because I'm kind of building off of your comments. And I'm trying to understand.

Because I agree with you, there's a

piece missing. But, I'm not sure I fully understand that piece that you're articulating.

DR. LORCH: And I'll just try to brief you something here.

MEMBER MOORE: Yes, please.

DR. LORCH: Hospital administrative data alone in our latest analysis is missing gestational age in approximately 50 percent of the premature access.

The crowd out of codes or with codes, so now you're starting to make up, this looks like it's a preemie. This looks like it's part of the denominator.

I would love it if we could just do it on hospital administrative data alone. But, from a validity perspective, face validity at least, we seem to need the linkage because the information in the birth certificate with specific birth weight and gestational age is very critical.

There are a few States that do provide that information in the hospital discharge

records. But, it's somewhat of a challenge.

The other thing with HCUP in particular, it's following infants through their hospital course. And identifying the discharging hospital which can be a difficulty unless you have linkage over the -- to collect it publically, I think with AHRQ they have some of the more -- with more identifiable to allow for that transfer link.

But, what we are missing explicitly there is the Vital Statistic data.

CO-CHAIR GREGORY: Matt?

MEMBER AUSTIN: Yes, so for me, and maybe I'm not quite fully understanding the data sources for this measure. But, my experience with readmission measures is that really health plans and maybe higher are where were can reliably measure readmissions.

Until we have a unique patient identifier, it's really hard to track readmissions across facilities. And so, that would be my concern from a reliability standpoint

with this measure.

CO-CHAIR GREGORY: Karen?

MEMBER AUSTIN: But I do think it's an important issue. So, it's not like I'm discounting the issue.

DR. LORCH: No. And to the point, I may have misunderstood, but the levels as well. So, again, I think it's whatever the group thinks is the, you know, the data we do present is at the -- it's using State administrative data.

Values are from individual hospital level. But, the data do come from a State level source. We did not choose to put in our insurance level source because it's again, from a liability it's potential issues with the data they currently have.

So, if that changes, kind of some of these, the framework of it that is, I mean, obviously that's totally okay. It's not -- there's not a specific, I'm not wedded to any of the potential levels. Just so the Committee to understand that.

And I'm going to have to go in a couple of minutes. I apologize to that.

CO-CHAIR GREGORY: Karen?

MEMBER SHEA: So no doubt I agree with all of the comments that this is an important issue. And one way in which we can look across the spectrum at the entire episode of care is to look at the insurer level data so that we can look at, you know, from birth through this time period, 30 days to see where the child's claims are coming from. If indeed they're coming from emergency rooms or, you know, different hospitals, et cetera.

There's two things though. One is that you're not presenting the measure at that level of evidence. And two, I noticed that there are multiple imputations that you mentioned for exclusions around perhaps I'm assuming diagnosis that you would expect maybe a readmission.

So, for example, a small premature baby who's got a planned readmission for eye surgery or for hernia repair. You know, how are

all of those planned readmissions excluded from the data set?

DR. LORCH: Using ICD-9 codes specifically. That's the best that we have from that particular element.

We did find that the number of those planned readmissions were small given current changes in practice. So, from that perspective.

And, I'm trying to think the imputation question for the --

MEMBER SHEA: So, within -- I didn't see them. But, you know, somewhere within this document, there's a list of all of those exclusions by ICD-9?

DR. LORCH: Correct. When we looked at those with the -- I'm sorry. Those are the congenital anomaly issues going back to the previous question that somebody asked about.

MEMBER SHEA: Okay.

DR. LORCH: So those are ICD-9 code exclusions.

CO-CHAIR GREGORY: Okay. Janet and

then Nancy and then we're going to move on.

MEMBER YOUNG: So, back to the hypothesis of this measure. Which is that readmission rates are an indicator of quality of care either at time of discharge in the outpatient setting, or with the provider who is taking care of that patient, or the patient's actual illness severity.

How does this particular data set help us drill down at the hospital level when there's very little the hospital can do to change those rates?

DR. LORCH: So, I think that gets into the big question about how -- whether that comment is actually true. I think that -- and many of us think that there is some element of readmissions that are likely not preventable. That's the noise in the measure.

But, many of the readmission for this age group are for conditions that one may argue are either discomfort on the part of the outpatient provider, or discomfort on the part of

the patient, whereby improved discharge teaching and improved transition of care from the inpatient to the outpatient setting, would reduce these readmissions.

Many of the readmission reasons that -- in this data set and others, other things like feeding, failure to thrive, neglect, along the opposite side, parental concern over a condition, which is a V Code.

So, I think -- and when -- I think I've said it, when I was off the line, I think the challenges that for every population and every case mix, it does differ what types of implementations you need. And process of treatment we may want to implement to minimize these readmissions.

I don't think -- and no one is actually saying that the rate should be zero. That's impossible. But, I think that what we do see is more hospital level, almost QI types of intervention to identify the root causes and their specific population.

That then may go onto reduce readmissions either at the hospital level. And I think it really depends on the system that one, is practicing into the larger system as well as the patients that we see.

We see -- and so, we have some data that's not published yet on that topic showing that people's discomfort with discharge and with families anxiety over going home.

Which is highly variable between patients. As well as variability on the outpatient side in terms of the transitions of care and the comfort of the patient.

So that's a kind of long answer to say I think it's more of a QI process, which is somewhat unsatisfying that we don't have a magic here's what we can do to get everybody to reduce readmissions.

I think like with other projects, I think with other readmission measures, I think it's really more on a QI framework at the local level to understand what the root drivers are for

such a variation.

CO-CHAIR GREGORY: Okay. Thank you. Final comment, Nancy. And then we're going to go back to reliability.

MEMBER LOWE: So I think my struggle with this is back in the specifications where the level of analysis is the facility or the State. And I don't -- that won't work in areas of the country like where I live.

Where our children's hospital serves the whole mountain west. And then we go over to where Kristi is in Salt Lake City. And they take the ones on the other side of the continental divide.

So, it's like the State is almost meaningless where we live. Because we care for -- our children's hospital, our catchment area is from the middle of Kansas, the Dakotas north, you know, Montana, Wyoming, Colorado.

And so, I'm really struggling with that. I totally agree. It's a great idea. But, I don't know how it will work.

CO-CHAIR GREGORY: Okay. I'm going to --

DR. LORCH: And I'm going to have to step away. I apologize for that.

MEMBER SHEA: I know you have to step away. But, I have one really quick question for you on process.

DR. LORCH: Sure.

MEMBER SHEA: And that is, did you make a distinction between actual admission and observation stay in your data?

DR. LORCH: No. Because of high variability between what a hospital may consider observation. We considered both of those a readmission. Just at the top.

Because in some places those would be considered a readmission. They don't have an obs unit. In other places, they're all on the obs unit for a certain period of time.

So, we made no distinction in this data set.

MEMBER SHEA: So, from your pulling

chart audit rather than claims-based data for that for that description?

DR. LORCH: No. I mean, no we didn't. It's claims. And I think what we -- what we have is just they were admitted to the hospital with these types of administrative data sets.

I think with the insurance you could. But I am a little worried about that because of pretty significant variability in what is considered an observation ad -- if you don't call it an admission, what an obs type of admission would look like between hospitals.

So we excluded that out.

CO-CHAIR GREGORY: Okay. Thank you. Good luck on your --

DR. LORCH: Okay. Thank you very much.

CO-CHAIR GREGORY: Moderating session.

DR. LORCH: Thank you.

CO-CHAIR GREGORY: Sheila?

MEMBER OWENS-COLLINS: I had a question for Nancy. I'm not sure since you have,

you know, you're one hospital for a large area, you would have a large data base.

So, I'm not sure, you know, why you would have problems getting readmission data?

MEMBER LOWE: But the children might be readmitted anyplace. Not in our hospital.

MEMBER OWENS-COLLINS: Oh, okay. All right.

MEMBER LOWE: So, like they go back to Montana.

MEMBER OWENS-COLLINS: Okay.

MEMBER LOWE: And if they have an acute illness in the next three weeks, they're going to be in a Montana hospital. They're not going to come back to us unless they need the level of care at which they'd get airlifted back to us.

So, --

MEMBER OWENS-COLLINS: Yes. But you could do a longitudinal, you know, study for the 30 days for all the patients that you've discharged and see what happens.

CO-CHAIR GREGORY: Sorry, I'm going to make one comment and then I think we should go back to the right way.

I know that using the California data set that he used, there's actually a record linkage number. And so you could in fact, it doesn't matter where you got readmitted. They could in fact identify those admissions at different hospitals.

So, I would like my discussants to summarize their feelings about the reliability. And then we will decide on whether we're going to vote or not.

I know, but I mean, if we need more discussion, we'll go with that.

MEMBER SHEA: So I believe we all have concerns about the ability to reliably collect the data which more so goes to feasibility. But, if you can't collect it then it's not going to be that reliable.

The other issue that I brought up last is the difference between a 24-hour stay or an

actual hospital admission. You're not going to see an observation showing up as an admission in the claims data set.

CO-CHAIR GREGORY: There's going to be a bill though.

MEMBER SHEA: I think, you know, the planned admission issues, you know, is another exclusion. It may be a few, but it's going to confound the data somewhat.

I think that, you know, looking at the ability to collect the data across the different states, New Jersey, New York, it really has to be at a payer level in order for this to make sense.

And I'm not seeing that the study has been done at the payer level. And so, I would say it's a good measure.

I mean, it's a good concept. It's something that's really important. But, I would say it needs to be baked more at maybe a payer level.

And just really show the evidence. And then come back to the Committee.

CO-CHAIR GREGORY: Okay. Cindy?

MEMBER PELLEGRINI: Yes, a question I think for Reva. And it may not be an entirely fair one. Because it's a very broad NQF question.

But, just looking into QPS, there are I'd say at least a dozen readmission measures. There's all cause readmission, there's readmission after myocardial infarction, after vascular procedures, after coronary intervention, after COPD hospitalization.

And I imagine that a lot of the things that we're talking about are not unique to this readmission measure. So, either are we holding this one to a different standard then other Committees have on other measures?

Or are there insights from some of those other measures that you might be able to give us on things like these patients don't come back to my facility?

DR. WINKLER: What I can tell you is that's a common question. But, I do think some

of you have raised something that is somewhat unique.

And that is the whole issue around children's hospitals and maternity hospitals in terms of the amount of difference between where patients go. So it's always a question.

I can't quantify it. But, it does seem that this maybe something that is somewhat different. And so, that would be the one thing I could say, Cindy.

MEMBER SHEA: I also think in this particular population where you are dependent upon the training and the competence of the parent to care for the child after discharge, it makes it a little bit different.

MEMBER AUSTIN: It maybe just worth looking at the level of analysis for those measures. I think they're mostly at the health plan level.

CO-CHAIR GREGORY: Sheila?

MEMBER OWENS-COLLINS: Yes, I agree, at the health plan. And also, the parent issue

that was just raised. That makes it even more important.

Because the parents are, you know, the primary people accountable. And it's very important that they understand.

And so, to the degree that they do or do not understand, it is reflected to how well we do with this measure.

MEMBER SHEA: Also, I want to bring up that the CMS readmission rate is for the same diagnosis within a certain period of time. And we're dealing with one admission, which would be for a diagnosis of prematurity.

And then the repeat admission could be for anything, including pneumonia, all cause, right.

CO-CHAIR GREGORY: All right. I am going to ask if we can vote on the reliability?

MS. ROBINSON-ECTOR: Voting is now open for reliability of measure 2893.

(Voting)

MS. ROBINSON-ECTOR: All the votes are

in. Four percent voted high. Twenty-six percent voted moderate. Sixty-three percent voted low. And seven percent voted insufficient.

So, for reliability the measure does not pass.

CO-CHAIR GREGORY: That failed. Okay. So, then we can stop?

DR. WINKLER: Yes.

CO-CHAIR GREGORY: You guys saved yourselves. Because now we can do Antenatal Steroids.

And without bias that should be a slam dunk. And you can go to lunch. Or no we can't. We've got eMeasures.

Okay. Antenatal Steroids. Okay. Developers?

MS. MILTON: All right. This is Celeste from the Joint Commission. This next measure is Antenatal Steroid initiation.

And what we're looking at in the denominator are those mothers that are delivering pre-term newborns at 24 to less than 34 weeks of

completed gestation. And of those, mothers in the numerator would be those where Antenatal Steroids were initiated prior to delivery.

The original measure was looking at a full course of Antenatal Steroid administration. And we learned that it was a problem for hospitals. It was a great burden to be looking for the second dose because in many times the mother had already delivered.

So, we did do a modification to where we're looking at just the initiation. Because we found that when they ordered it, they always ordered it and said repeat it in 24 hours.

And the goal here of course, is to get this at 100 percent.

CO-CHAIR GREGORY: Discussant?

MEMBER BELL: Hey, this is Amy Bell. There has been no new evidence since the last time this measure was endorsed.

And so, we actually move that we accept this based on past evidence.

CO-CHAIR GREGORY: Any opposition to

that?

(No response)

CO-CHAIR GREGORY: Let's talk about the gaps.

MEMBER BELL: Do you all have anything else to add about the gaps? No? Okay.

Although there has been improvement over the last four years, the last data being in 2014, there still is a significant gap in the performance. So, we recommend that we continue to endorse this measure based on what has been shown with opportunity for improvement.

CO-CHAIR SAKALA: And just to clarify, pretty impressive movement in three years, 54 percent to 82 percent. And again, a very wide practice variation around that.

CO-CHAIR GREGORY: I'm going to call for a vote.

MS. ROBINSON-ECTOR: Voting is now open for performance gap of Measure 0 -- oh, I'm sorry.

CO-CHAIR GREGORY: Question?

MEMBER KEATS: Yes. In terms of new evidence, ACOG put out the practice advisory a month ago about potentially extending this now to 36 and 6/7th rather than 34 based on something called the Antenatal Late Preterm Steroid Trial.

So, is that going to going to be incorporated at some point into this measure? Or how does that affect what's going on here in terms of evidence?

DR. MAIN: That is brand-new news. That paper was presented and published in February of this year.

So, that was certainly long after this was submitted. But, I think we want to be sure that this becomes a part of routine practice.

We're not sure that every single 36-week mom should get steroids if they come in in normal spontaneous labor at that point. So, the exclusions are still to be determined in that population.

It's pretty rock solid under 34 weeks. And that's where we intend to stay for the time

being.

MEMBER GOYERT: So the corollary would be what about going the other way? You're going to leave the measure as it is, but I'm talking about the 23 to 24-week window.

DR. MAIN: That does vary from center to center. I think 23 weeks would be a reasonable choice. There are mothers who decline intensive care at 23 weeks.

And I think we'd have to have an exclusion for that if we were going to go there. Again, I think the intention was to say what is the rock solid areas that everybody agrees on and is focused on those.

CO-CHAIR GREGORY: Sindhu?

MEMBER SRINIVAS: At what point, I mean this is showing great improvement over sort of in trending the time. And what point do we think about like when do you actually retire a measure?

Or is that more a philosophical question I guess? I'm just more -- I'm asking

kind of for it.

DR. WINKLER: Actually, that's sort of the question on the table right now under opportunity for improvement. Because really, the underlying question is, what's the quality problem?

What does the data show as the quality problem? And so, that's what you're evaluating on the opportunity for improvement, is whether there is continued improvement. And there are a lot of factors that go into it.

CO-CHAIR GREGORY: Dr. Bailit?

MEMBER BAILIT: I know when we looked at this in the OPQC, Ohio Perinatal Quality Collaborative Network, what were found improvements were from was increased in better coding. And that was still more administration.

For the more -- when people were getting up into the high 90s, they were like, oh, the baby's crowning. Here's your shot. Okay. You can push.

Like there's got to be a it's too late

part. And you don't want people doing that just to get their box checked off.

So, to the extent that I think -- I think was Sindhu, was it you who brought that up? I think there's a real point there.

The numbers would suggest that we're not as good as we could be. But, I don't know how much of that is coding versus how much of that is administration.

MEMBER BELL: I would just add one thing about delivery within a few hours of actually the patient being admitted. If there is a qual that the provider rights in the notes that addresses that, then that patient is actually excluded from the measure.

MEMBER BAILIT: That depends on the center. In other words, some places are giving the shots up until literally crowning.

MEMBER BELL: They don't --

MEMBER BAILIT: It's ridiculous.

DR. WINKLER: Would you call that an unintended consequence?

DR. MAIN: On the other hand, if someone is six or seven centimeters and you think they may deliver, they may not for even a day or two.

So, you know, of course crowning is silly. But six or seven centimeters is not necessarily a bad move.

CO-CHAIR GREGORY: And there is some data to suggest that we're giving too much. And that you know, we really want them to deliver it within 48 hours to seven days. And that came out also in February.

But, we are voting on the measure in front of us right now. And so, I think that based on the information we have, do we think that there is a gap and still an opportunity for improvement.

MS. ROBINSON-ECTOR: Voting is now open for Measure 0476 for performance gap.

(Voting)

CO-CHAIR GREGORY: Someone left.

MS. ROBINSON-ECTOR: Yes. All the

votes are in. And voting is now closed. Twenty-seven percent voted high. Fifty-four percent voted moderate. Nineteen percent voted low. And zero voted insufficient.

So, for performance gap of Measure 0476, the measure passes.

CO-CHAIR GREGORY: Okay. So, validity. I'm sorry, reliability. I think we can take previous. Is there any new reliability testing I need to know about?

MEMBER BELL: No.

CO-CHAIR GREGORY: And the same for validity? Oops, Jaleel, I'm sorry.

MEMBER MAMBARAMBATH: I have a question again about the denominator exclusions. One is enrolled in clinical trials. So, that's something that should be brought up again.

These are high-risk moms who will most probably in academic centers be enrolled in one trial or the other. And it's probably prudent to take that out.

The second one is more a part for me

from my point of view is the documented reason for not administering antenatal steroids. So, that is giving an out for the physician who's taking care of the patient.

And what I can tell you from my own institution, two of the documented reasons for not giving steroids was preeclampsia or any pregnancy induced hypertension, and diabetes. That constitutes about 50 to 60 percent of the population.

So, on paper they have 100 percent concurrence with the measure. But when you look at it, the real numbers, it was very low. And it took a lot of time to convince them to change that practice.

So, it would be good to take it out as well.

CO-CHAIR GREGORY: Okay. Juliet?

MEMBER NEVINS: Or just be more specific as to what is an appropriate exclusion.

CO-CHAIR GREGORY: Go ahead.

MS. MILTON: I wanted to make sure.

There are certain things that would be considered appropriate exclusions like chorio -- I can't ever say it, chorio. You all know what it is? Okay.

So, you wouldn't, if that was present, that would be a reason why you wouldn't do that. We also look at they're saying it's an imminent delivery.

And we've talked without technical advisory panel and they suggested that within the first two hours you're going to know whether that's imminent or not. So, you should be initiating that dose.

If they deliver six or seven or eight hours later and didn't get that, that wouldn't be considered an imminent delivery. Therefore, it wouldn't be a reason for not initiating it.

Also, we would look at any case where they know that the fetus has anomalies that are incompatible with life. Again, this would be a reason why you wouldn't initiate steroids.

MEMBER MAMBARAMBATH: So, there is a

risk.

MS. MILTON: We provide notes for abstraction to the abstractors to be looking for this sort of documentation. And I'm kind of surprised about the diabetes thing, because that's never ever been a question.

And I want an answer to all the questions.

MEMBER MAMBARAMBATH: So, there are concerns within the OB community, at least in the place where work. And I'm assuming that there are many other places too. At least some other places too.

That the initial studies which were done with Antenatal -- initial studies which were done with Antenatal Steroids did not specifically include the moms with hypertension and moms with diabetes.

Or -- so, and there is a significant concern in the OB community that well, if you are going steroids to these moms with hypertension, are you going to delay the delivery of that

patient for the steroids to act?

So, with the hypertension, are you putting the mom in more danger? And are you putting the fetus in more danger?

So, that is a concern that they had. Or at least our OB community had. And of course a significant population.

CO-CHAIR GREGORY: I think -- all right, I'll be blunt. We would be more than happy to come give grand rounds there. Parkland's unique.

Greg? Nancy? Okay. All right, so -- go ahead Rajan?

MEMBER WADHAWAN: Just a clarification question for you also. When you say the reason for not initiating Antenatal Steroid therapy, is it sort of a free test choice or is it only a few categories that you can pick?

Because if that is being the case, that preeclampsia and diabetes shouldn't even be there. That shouldn't even be allowed.

And some of the other ones that are

valid like somebody's crowning and shows up at 24 weeks and delivers should be included.

DR. MAIN: The three on the list that I recall are chorioamnionitis, imminent delivery, and anomalies and compatible with life.

CO-CHAIR GREGORY: You want one more? Go for it.

MEMBER MAMBARAMBATH: Yes. The practice has changed now at the hospital, so.

(Laughter)

CO-CHAIR GREGORY: All righty then.

MEMBER MAMBARAMBATH: But I think it's an important question. Because if you are putting that as a documented reason, and if it is not only these three or four, because I remember from my -- from previous data that I have looked at from my own institution, it was under question for those patients who they felt that it was not manageable.

CO-CHAIR GREGORY: Okay. We are going to vote on the reliability of this measure.

MS. ROBINSON-ECTOR: Voting is now

open for reliability of Measure 0476.

(Voting)

MEMBER MAMBARAMBATH: Are we taking out the enrolled in clinical trials in this one? Or no?

MS. MILTON: Yes, we are. We're removing that.

CO-CHAIR GREGORY: Do the discussants have any comments regarding validity? Oh, I'm sorry. I'm trying to catch us up. I am so sorry.

MS. ROBINSON-ECTOR: Okay. All the votes are in. And voting is now closed. Eighty-eight percent voted moderate. Twelve percent voted low. And zero voted insufficient.

So, for reliability of testing for Measure 0476, the measure passes.

MEMBER BELL: We do not have any additional comments related to validity. I mean, empirical validity testing did occur. So, we recommended it to continue.

CO-CHAIR SAKALA: Okay. I could just

note a few changes in this. Not -- there's the continuous feedback and collection.

They switched to ICD-10 codes. The numerator changed from steroids administered to initiated. And the denominator broadened from it used to be 24 through 31 weeks. Now it's through 33 weeks.

I think those are the important changes in their measure specification.

CO-CHAIR GREGORY: So, shall we vote on the validity? Yes? Vote on validity. Go for it.

MS. ROBINSON-ECTOR: Voting is now open on validity for Measure 0476.

(Voting)

MS. ROBINSON-ECTOR: All the votes are in. And voting is now closed. Fifty-six percent voted high. Forty-four voted moderate. Zero voted low. And zero voted insufficient.

So, for validity of Measure 0476, the measure passes.

CO-CHAIR GREGORY: Feasibility.

MEMBER BELL: So, for feasibility, facilities are currently collecting this data. As the facility number of deliveries drops from 11 hundred to three hundred, it may become more burdensome for those smaller facilities to collect that data.

Although a lot of those facilities would not be delivering patients that are less than 34 weeks to start with. But, if there's, you know, inclement weather, things like that, they're going to have some of those here and there.

But, feasibility we didn't see any issues really with this either.

CO-CHAIR GREGORY: So, shall we vote on feasibility?

MS. ROBINSON-ECTOR: Voting is now open for feasibility of Measure 0476.

(Voting)

MS. ROBINSON-ECTOR: All the votes are in. And voting is now closed. Sixty percent voted high. Forty percent voted moderate. Zero

voted low. And zero voted insufficient.

So, for feasibility of Measure 0476, the measure passes.

CO-CHAIR GREGORY: And finally usability and use.

MEMBER BELL: This measure is currently publically reported. And is used in accountability programs. We do not see any new information.

CO-CHAIR SAKALA: I could share that this is -- as a population measure, is included in the MAP Medicaid Adult Core Set.

CO-CHAIR GREGORY: So, we shall vote.

MS. ROBINSON-ECTOR: Voting is now open for usability and use of Measure 0476.

(Voting)

MS. ROBINSON-ECTOR: All votes are in. And voting is now closed. Eighty-five percent voted high. Fifteen percent voted moderate. Zero voted low. And zero voted insufficient.

So for usability and use of Measure 0476, the measure passes.

CO-CHAIR GREGORY: Okay. So our next measure is Measure -- oops, I'm sorry.

(Laughter)

CO-CHAIR GREGORY: Shall we endorse this Measure? Shall we vote to recommend for endorsement of this Measure?

MS. ROBINSON-ECTOR: Voting is now open for overall suitability for continued endorsement of Measure 0476. Option One is yes, and Two is no.

(Voting)

CO-CHAIR GREGORY: We need to put that on our script. That extra little step. That's what it is.

MS. ROBINSON-ECTOR: All the votes are in. And voting is now closed. One hundred percent votes yes. And zero voted no.

So, for recommendation for continued endorsement of Measure 0476, the measure passes.

CO-CHAIR GREGORY: So, we have five minutes before we do our public comment. Okay, we'll do it early.

So this is my suggestion. My suggestion is that we break for lunch, do a working lunch to get through the rest of these measures.

Alternatively, we can power through the next measure and then break for lunch. So, how many people want to do a working lunch?

(Show of hands)

CO-CHAIR GREGORY: Perfect. Okay, operator, we would like to open the line for public comments.

OPERATOR: Okay. At this time, if you would like to make a comment, please press star then the number one.

And there are no public comments at this time.

CO-CHAIR GREGORY: Okay. Is there anyone in the room who would like to make a public comment? Perfect. Thank you, sir.

DR. SAN ROMAN: Hi. This is Dr. Gustavo San Roman again. And unfortunately I wear two hats at this meeting today.

One is as a developer. And one is as a member of the public. And it is Committees like this one that are given the responsibility to provide guidance for other stakeholders that are perhaps not as skilled at math or science.

Unfortunately, without scientific guidance, other stakeholder may move forward with plans that lack math or science. As is now evidence in Section 5.03 of the new 2017 contract for Cover California.

The new contract states that contractors must exclude hospitals from their provider networks if the hospital is unable to achieve an unadjusted NTSV C-section rate below 23.9 percent.

Twenty years of science and math confirm the importance of risk adjusting the NTSV Cesarean birth rate. If that rate is to be used as a Cesarean birth measure. And until this year, Measure 0471 has had a risk adjustment for age.

Just to be clear, this Committee has

just endorsed the removal of the risk adjustment from that Measure based on only one graph with data only from California that claims that the risk of age is completely cancelled out by BMI.

I am dumbfounded that anyone can come to the conclusion that age and BMI completely cancel each other out based on one graph and one published study when the published study provided doesn't even include BMI as a risk factor.

Extensive prior research has shown that the NTSV Cesarean birth rate increases with age within a hospital.

Since age and BMI do not cancel each other out within a hospital, in order to claim that age and BMI cancel each other out when comparing hospitals, somehow every hospital across the nation must attract the exact combination of NTSV patients where a risk due to age and the risk due to BMI completely cancel each other out while still maintaining a rate that increases with age.

Math and science do not support the

claim that age and BMI cancel each other out. Every hospital across the country with older NTSV patients will be adversely affected by a Cesarean birth measure that does not have any risk adjustment for age.

This adverse effect will soon become evident to hospitals in California. More concerning is that hospitals with unadjusted NTSV Cesarean birth rates of over 23.9 percent will not be able to justify their higher rate by claiming that they have higher risk patients.

This is because according to the Joint Commission, a risk adjusted model, which included age, BMI, race, hypertension, and diabetes, found differences limited to only one to two percent.

Or in other words, all hospitals should be able to achieve an unadjusted NTSV Cesarean birth rate within one to two percent of the average regardless of their patient population.

Six years ago, I was the only one who recognized the flaw in the risk adjustment for

Measure 0471. And today, I am the only one who recognizes the disaster that is coming with an unadjusted NTSV rate.

To fully understand the disaster that is coming to California, one only needs to look back to 1997. That was the year with the lowest reported national unadjusted NTSV Cesarean birth rate.

And that year saw an overall Cesarean birth rate of 20.8 percent. The date from 1997 revealed that in 1997 if a hospital had NTSV patients with an average age of 34, they would have had an unadjusted NTSV Cesarean birth rate of about 27 percent.

This means that if a hospital in California that has 34 year old NTSV patients currently achieves a rate as good as the best year on record, it will still be excluded from the Cover California provider network.

God help those hospitals in California. And God help us all if this Committee doesn't check the math. Because a

flawed Cesarean birth measure is worse then having no measure at all.

Thank you.

CO-CHAIR GREGORY: Thank you. Are there any other comments?

(No response)

CO-CHAIR GREGORY: All right. Then I'm going to charge the Committee to get lunch. And then we'd like to resume at 12:30.

(Whereupon, the above-entitled matter went off the record at 12:14 p.m. and resumed at 12:29 p.m.)

CO-CHAIR GREGORY: I'm very proud of you guys all coming back on time. If we have our developers at the table, could I have everyone's attention?

We're about to get started again. Actually before we do the next measure, we're going to get a primer on eMeasures so we can chew our food.

DR. WINKLER: All right, eMeasures. Over the last few years, there's been a real

growth in the development of eMeasures, and just to be clear, what we're talking about are not just measures that use, say, electronic health records as a data source.

But eMeasures, or the federal government calls them eCQMs, electronic clinical quality measures, are in a very specific format. The HQMF, health care quality measures format, that is an industry standard.

And so we are talking about a measure that is specified using that HQMF format. It also specifies value sets that are registered in the National Library of Medicine at the Value Set Authority Center.

So an attempt that as eMeasures are being developed to put some structure around the idea of an eMeasure. Okay?

So it's not just if you're using EHR data and, because you can go into an EHR and abstract all sorts of data, but that's not an eMeasure. All right?

So I do want to make sure we

understand the difference. Now how can eMeasures get developed? Well, all sorts of ways we're finding. You can make a brand new eMeasure.

And one of the hopes and dreams about eMeasures is that with the capabilities of EHRs that we don't have in other data systems, we might be able to start seeing new and greater and better and more wonderful measures, still waiting. But nonetheless, that's, those are the hopes and dreams.

And so that's a de novo measure. I mean, it's not going to have an antecedent other type of measure. It's its own thing. We're starting to see an occasional one of those, but not that many.

What we're tending to see are measures that are based on existing measures. And you know, it's totally internal to NQF and probably irrelevant to you all, you know, whether we call them respecified or legacy because they're in use in federal programs. Doesn't matter.

The point is there is an antecedent

existing measure based either on claims, paper, registry, something. And the eMeasure is an attempt to create the -- or the eMeasure is a new version.

Because of the unique aspect of eMeasures, NQF treats them as separate measures, and they have numbers that are different, although we are in the process of reworking our data system to combine the numbers of the two versions of the measures so that you know they're two versions and they are kind of all related to the same parent measure, but yet are distinctly different.

So what you see in our next group of measures is two of the Joint Commission core measures and their new eMeasure version. And so go to the next slide.

And so technically I'll tell you these are what are known as legacy measures just because they are in use as the original version, and the eMeasure is now the new version, or another version, if you will.

And so you can start to see how the numbers are going to get a little confounded, but the eMeasure version will relate back to the original antecedent number.

And so what we do is we will look at them as one and then the other because there's a great deal of crossover. And you'll find in the measure information form that essentially the information around evidence and opportunity for improvement is an identical, you know, information on the two, so re-having that conversation twice doesn't make a whole lot of sense.

So we'll be able to hopefully look at some things a little bit more efficiently. One of the things about eMeasures is the challenge that developers are struggling with in terms of being able to formally test those measures for reliability and validity because a lot of the use of eMeasures, particularly through the federal meaningful use programs, has relied on attestation of the provider to say I can do it.

But there's no data. So in the absence of numbers, there's really hard to do any analysis on numbers. So this is the quandary that is evolving and resolving over time, and we can anticipate seeing more data, you know, fairly soon actually.

But nonetheless, we're fairly limited now. As an alternative to support the, you know, the understanding about the data reliability and validity, if you will, and feasibility, there is, there are synthetic tools that can be used where you create a database.

You know, it isn't real patients, but you create a data set, and then test the measure against that data set to see if you can produce results, and whether those results were, reflect, you know, what it is you input.

So that is acceptable at this point in time. This is a transitional kind of thing. I mean, it's going to change as we are able to collect data.

I know one of the questions we asked

on the workgroup calls that the Joint Commission was able to help us with was understanding that now the Joint Commission's collecting data with the eMeasures and there are a certain number of healthcare organizations submitting data actually through the eMeasure, as well as the traditional avenues.

So that's the world of eMeasures at this point in time. And so, does anybody have any questions? Okay.

PARTICIPANT: So can you just kind of, a little bit more about the, there's a fake data set?

DR. WINKLER: Right. There is a fake data set. Okay, there are several tools, but the most commonly used one is called Bonnie, and it is, you know, you create a synthetic data set. You know --

PARTICIPANT: It has discharge codes and --

DR. WINKLER: It has the codes you need and the data you need to make it work.

PARTICIPANT: Is that commercially available?

DR. WINKLER: Yes. Actually it is commercially available free of charge is my understanding. Yes. It's a fed, the feds created it. Diana.

MEMBER RAMOS: Has the accuracy of the data that's been inputted been analyzed, because obviously if it's all going to be free, you want to make sure that the information is correct, accurate, so that the interpretation can be reliable.

DR. WINKLER: I think what we need, one of the conversations that's probably reasonable to have is with the folks, because the Joint Commission isn't, is doing this measure with eMeasures, although, you know, still early, not tons of experience, but you've got some.

And those are the questions we'll ask them. You know, what are you doing in terms of looking at your data. Okay? Any other questions about eMeasures in general before we get to the

specifics of these measures?

PARTICIPANT: May I ask a process question? So when they review the measures are they reviewing them as independent of each other?

DR. WINKLER: Yes. However, we'll be able to carry over a lot of -- any of the similarities, so we don't have to have the same conversation, you know, when it doesn't apply. But you will be looking at them twice, or separately. Okay. With that, back to the chairs.

CO-CHAIR GREGORY: Sorry. We're going to do measure 0469. No. Is it? Yes. Elective Delivery by the Joint Commission as the developer, and our discussants are Tracy, Sheila, and Jennifer.

MS. MILTON: Hi. Celeste from Joint Commission. Now, what we're looking at and the denominator here are going to be those patients that have delivered at 37 weeks of completed gestation to less than 39 weeks, so it would be 38 and six days, and of those patients, how many

had an elective delivery.

The goal here is to reduce that number, but I think we've already had a little bit of conversation that the goal is not to get to zero.

It has never been the intent, because there are going to be certain circumstances that are virtually impossible to identify all of.

One example would be a patient that needs to receive chemotherapy for a certain type of cancer. That may not hold true for every patient that has cancer. It could be a specific type.

So there's going to be circumstances where some patients will need to be delivered early, and there is a medical indication.

But what we've done is we've tried to get the majority of those, and our analysis has shown that we've gotten about 98 percent of them by the use of ICD codes and also with the data on it called prior uterine surgery, where specific types of prior surgeries would be reasons to

perform an early Cesarean delivery. And that's it.

CO-CHAIR GREGORY: Discussants?

MEMBER FLANAGAN: So just to say a little bit more about this measure. I think for many of us in the room, we're very familiar with this measure.

It was, as mentioned by the presenters, that it was meant to decrease non-medical elective deliveries that were pre 39 weeks. And the intent was to decrease neonatal morbidity and mortality.

Also with the hope of decreasing the rate of C-section, which we'll get to a little bit later. It was originally endorsed in 2008, and the most recent endorsement was 2012. It is a process measure and I think I'll stop there as far as introduction to the measure.

CO-CHAIR GREGORY: Is there any new evidence that you'd like to share?

MEMBER FLANAGAN: The evidence, should I keep going?

CO-CHAIR GREGORY: Yes, please.

MEMBER FLANAGAN: So the evidence put forward was last discussed in 2012. I don't see any additional evidence.

And, but just as an update, ACOG reaffirmed the practice bulletin for induction of labor, including information in this area, and there was an ACOG opinion that reaffirmed the evidence for non-medically, for, and provided guidance for non-medically indicated early term deliveries.

CO-CHAIR GREGORY: So I would like, unless there are any --

MEMBER BAILIT: I actually, I actually do want to add some new evidence that wasn't listed but I think is relevant.

And this is actually supporting of the measure, which is, there's always been a concern that were we increasing the stillbirth rate with the NTSV going down, and I think there is new evidence that we are not causing harm of the stillbirth rate.

So to the extent that that's not directly to the point, but it is, I think, additional confirmation of safety. It's relevant.

CO-CHAIR GREGORY: So unless anyone is opposed, I would like to suggest that we vote to accept this evidence.

MEMBER OWENS-COLLINS: I just had a question I wanted to clarify.

CO-CHAIR GREGORY: Please.

MEMBER OWENS-COLLINS: Okay. And I agree that the concerns about not doing C-sections have been, or have not been founded. But I do have a question about the terminology when you talk about elective versus medically necessary versus medically indicated.

I mean, there seems to be some contraindication there. Because when I first read it, I was, my first thought, and I asked one of my OB colleagues, I mean, who would do an elective C-section. I mean, elective is just, I mean, it's just, for a neonatologist, I guess --

CO-CHAIR GREGORY: Who will address? Yes, so who would do that?

MEMBER OWENS-COLLINS: Right. And would write and document it secondly. So, right, so, right and document it is the biggest part. So you know, I think that, you know, so the example of someone that has cancer, I wouldn't necessarily consider that elective.

I mean, that's maybe more medical necessity or medically indicated for the mother. Now I think we need to be careful with those terms because, you know, I mean, elective is going to raise flags and potentially, you know, get people in a lot of trouble. So it's really medically indicated.

MEMBER FLANAGAN: So let me add that there is a huge appendix that spells out pretty much every medical condition you can think of, but it's never going to be completely inclusive. There's always weird things that come up.

MEMBER OWENS-COLLINS: Yes.

MEMBER FLANAGAN: There are medical or

real indications, but there's a huge list of what's considered medically indicated.

CO-CHAIR GREGORY: Juliet?

MEMBER NEVINS: So I just wanted to comment and to answer your question. There are many obstetricians who will do an elective Cesarean section. You know, the patient is afraid of pain, she demands it.

There are many obstetricians who feel that the woman has a right to choose her mode of delivery. So it's, you know, I've seen it done many times.

CO-CHAIR GREGORY: Okay. Without any further comments, I suggest that we vote on the evidence, and that we vote to accept the evidence that was previously, oh, you're right. You're right. That's right.

Okay, so that means that we can now go to reliability. I'm sorry, the gap. I'm tired. Let me get my little script here.

MEMBER FLANAGAN: I'll talk faster. So there still remains a gap, although the gap

has been narrowing over time. The rate has gone from 13.6 down to 3.3.

However there is going to be an additional number of hospitals coming onboard in January 2016, which should had, which should include 80 percent of all birthing hospitals.

So I think we would expect to see the newly reporting hospitals to show wider variation than we're seeing right now. Right now the variation is between zero and 8.7 percent, roughly. Between 10th percentile and 90th percentile there's of course a wider range than that.

But the variation has narrowed, but it will probably go up again with the inclusion of new hospitals. And by the way, there's a typo on this document.

It says all hospitals with greater than 110 births. I think it meant to say 1,000 and 110 births. And it's consistent through the document, that typo.

CO-CHAIR GREGORY: Okay. Jennifer?

MEMBER BAILIT: So in terms of the evidence gap, I am concerned with this one in particular because of some work that I did prior, while at HHS, and I also know that this work is now going to be published.

And we brought this up during our workgroup call and we talked about this issue of gaming of the system and then, you know, this notion of the definition and recognizing that we're trying to capture early term elective deliveries that are non-medically indicated.

There's a lot of subjective terms in that, which makes, which makes this challenging. And I'm not saying that we shouldn't collect this information, but I think that we need to recognize the limitations of this, and also be cautious.

I know there was some sidebar discussion about, should we retire this measure because we've met our goals.

But I just think that we really need to think about what is it that we're measuring,

how are we measuring it, are we really capturing it, and then bringing into light this issue of gaming of the system which we all know exists and is happening.

And if we really want to move the needle on this issue, how do we, how do we measure this topic in a way that's meaningful, and we don't assume that we've accomplished our goals.

CO-CHAIR GREGORY: Do you want to --

MEMBER FLANAGAN: I think for those of us who have been looking at this measure for a long time, I'm just going to respond to what you said. I think one of the glaring things that comes out are non-medically inductions that happen after 39 weeks, which is not before us, but I think that sort of speaks to some of the issues you're talking about.

CO-CHAIR GREGORY: Actually I'll speak to it too. I think that what happens is you get so good at it, but then you forget. And as soon as you forget, it pops back up again. So, in

fact, they wait for you to stop looking.

DR. MAIN: I think the issues around this, around the indications that are covered in this measure can, are in several different categories.

The most common one that was a problem before was scheduling of elective repeat C-sections. Not primary C-sections so much, but repeats.

And that's where the actually most of the neonatal morbidity was, was doing a section at 37 weeks. That's pretty much eliminated now. You know, and I think that's a big change.

I think some of the induction indications are being shaved a little bit, but that's not the big driver of morbidity here in my book.

So you know, you can't police every last birth and every last hospital, but you're looking for big trends, and big numbers, and I think we're actually making a big impact.

MEMBER MOORE: So I guess my question

back is, should we be, and I've sat in meetings with ACOG on those where the question is, do we, do we capture elective induction of labor as the measure, or do we look at overall induction of labor?

Because if we are reducing the elective, then we're overall reducing induction instead of trying to deal with this gray zone and trying to deal with some of the inherent issues with trying to capture elective. And I'm not expecting you to have an answer.

DR. MAIN: That's why there's a pretty long list of ICD codes that makes a stab at it. They're pretty loose though. It's any hypertension, any diabetes, which you know, gestational diabetes really shouldn't be delivered at 38, 37, 39 weeks.

So it's looser than what many people would like it to be. But it's also trying to balance out, you know, what, you know, working with obstetricians in the field to --

MEMBER MOORE: I would never advocate

for using ICD-10 codes, so that's an area where because of the way that the payment system is for maternity care, typically we lose, in terms of capturing those other data sets, we have pretty fair accuracy in terms of vaginal birth and C-section, but the other subsequent measure, or codes aren't as accurate. So I would agree with your statement.

CO-CHAIR GREGORY: Sheila.

MEMBER OWENS-COLLINS: I don't really have anything else to add. I think your limitations are, have been well stated before, as well as the pros of this measure, so --

CO-CHAIR GREGORY: Sindhu? Your card is up, that's why I called on you.

MEMBER OWENS-COLLINS: Oh, got to fix it.

MEMBER SRINIVAS: I just wanted to say something and comment to the issue of all inductions. I think while the indications are loose and lead to some area for a little bit of fudging here and there, I, in practice, have also

seen the sort of opposite happen where sick women with preeclampsia don't get delivered.

And while the stillbirth rate might not be sort of increasing, there's a lot of, you know, other balancing sort of maternal morbidity and other things that I think anecdotally lots of providers around the country have seen and sort of pushed back, like oh, I can't deliver her because she's not 39 weeks, but she actually does have a medical indication.

And so there's a little bit of adverse creep as well. So when you start talking about all inductions, I would really worry about that actually having a negative impact on maternal health.

CO-CHAIR GREGORY: Cindy?

MEMBER PELLEGRINI: So at the March of Dimes, I get to hang out on a daily basis with a whole lot of people who think this is still a really, really important measure.

This is brand new practice change, right? We're talking about this rate going down

over three years, and I think Dr. Gregory is absolutely right that that's not enough time to just say, okay, our work is done, and walk away. Because it will simply come right back up.

We need to give this more time to be in place. But it's also continuing to play, I think, an important role in a couple of other areas, which may not by themselves be reasons to maintain a topped out measure, but in addition to everything else illustrated, it's important.

One of them being that, is that this measure has played and continues to play a very important role with policy makers as kind of their entry point into understanding and learning about quality improvement.

This is kind of showcased as one of the, as an understandable example of how we can drive quality change, quality improvement, and effect, both cost and outcomes.

It's something that we talk about a lot on the state level when we're just trying to start explaining to an elected official what this

whole quality improvement enterprise is.

I think it's also continuing to play a role in understanding the processes, the sort of different stages that a measure goes through to full success.

So right now we're happily in the stage of, how do we get those outliers? How do we get those late adopters that are still not at the rate we want them to?

What are the additional barriers and how do we communicate or support them? So just a strong vote to continue, to keep this in place.

CO-CHAIR GREGORY: Deb?

MEMBER KILDAY: And I'm just going to jump on your bandwagon. Having the opportunity to hit some of those outliers, I go to those hospitals. I work with them.

I am often amazed at some of the gaming that goes on, even in those that are meeting the measures, and you do sort of that quality improvement work at the local level.

So I see both ends. I see a lot of

gaming. I'm just going to put it out there. It's happening. But also see, I see primarily smaller institutions where there are cultural barriers, and they are really struggling with implementing this.

And I also hear from organizations that have implemented it, their provider's feedback in that they can't wait for it to really relax because they want to go back to the old way. And don't think for a minute that isn't there.

CO-CHAIR GREGORY: Matt.

MEMBER AUSTIN: So to continue the bandwagon, I would reinforce the notion that we still have opportunities with this measure.

The Leapfrog Group has actually been measuring hospital performance on this for now seven years, since we've had the opportunity to look at hospital data longitudinally on this measure.

There's still a fair number of hospitals that remain with significant rates, and

we actually see some hospitals who are moving in the, quote, wrong direction, i.e. they are -- had low rates and actually are climbing back up. So I would vote for continuing to put importance on this measure.

CO-CHAIR GREGORY: Diana?

MEMBER RAMOS: Yes, in Los Angeles County where we have 60 delivery hospitals, we actually are bringing down the state in terms of, well, we have the highest rate of primary C-sections, and just, you know, just we need a lot of help on the quality measures.

So I think this is really an opportunity, even though California as a whole looks good, Southern California is the one that's bringing down the state.

CO-CHAIR SAKALA: So just a comment, I did the calculation for 2014, looking at those two weeks gestational age, and 3.3 percent performance, it's under one percent of all of our babies that then fall into that still vulnerable range and recognizing that it should never get to

zero.

And looking at most of those babies who will do well, I'm not arguing for retiring this, but I just want it to fit into the discussion about, you know, big time impact, and point out, and I also agree that I think we need to wait and see what those hospitals look like with 300 or more births, the point you made, Tracy. But I just want to point out what the actual numbers are there.

CO-CHAIR GREGORY: One more.

DR. MAIN: One quick point about the potential for gaming. I think everybody recognizes there are some cases here and there, but I think sort of the proof of the pudding is the reports first from Ohio and then from the National Center for Health Statistics that the rate of 38 and 37 week births in the United States has fallen significantly.

And that's sort of the bottom line, if you would. I think there are other collateral benefits of this, which is that the rate of 36

week births has fallen significantly.

You know, and lowering the pre-term birth rate in the United States, as people start reevaluating who needs to be induced and when. So I think those are two nice benefits overall.

CO-CHAIR GREGORY: Last comment. Naomi?

MEMBER SCHAPIRO: I'll be, I'll be brief. I just wanted to emphasize the importance on the pediatric end of this in that it was really only recently, if you think about the long term, that we have to even recognize that 37 to 39 weekers have significantly more problems and more hospital re-admissions.

And I don't think that's really penetrated to the entire like population of moms. So I think the more we can keep this measure in the public eye, that's really good for kids.

CO-CHAIR GREGORY: So if no one is opposed, I'm going to call for a vote on the, whether or not there is a gap in opportunity for improvement.

MS. ROBINSON-ECTOR: Voting is now open for performance gap of measure 0469. Okay. All the votes are in and voting is now closed.

Forty percent voted high, 52 percent voted moderate, eight percent voted low, and zero voted insufficient. So for performance gap of measure 0469, the measure passes.

CO-CHAIR GREGORY: Okay. Discussants, can you share with us any new information about reliability?

MEMBER FLANAGAN: So reliability, a couple things that have happened in that there have been changes through the years to actually more specifically measure what I think it was intended to measure.

Just to lay those out, first of all, there's going to be, this is not in the past just coming, the ICD-9 being converted to ICD-10, and it's unclear how that's going to affect the reliability of this measure.

The numerator also included population, the numerator included population for

medical induction of labor now requires a check for the presence of labor prior to the procedure of induction.

And that was, I think, and enhancement actually to the measure. Also, prior uterine surgery was better defined, and there was also some work around, I believe it's excluding the -- unable to determine no prenatal care patient who comes in where you can't determine the gestational age and that was considered a ding.

I think the one thing that many people have talked about that still troubles, especially small hospitals, is that sampling is allowed. And because sampling is allowed, and the specifications for sampling, it allows for wide variation from quarter to quarter because we have such low numerators.

I can speak personally from my own system that you can go wildly, you know, you happen to be unlucky one time and your sample of 25 includes two outliers, you can go wildly between zero and, say, even 12.

So that's the one criticism, and I think the -- when there's no sampling, and there's larger numbers, that variation goes away. I'll stop there. Anybody want to add anything?

CO-CHAIR GREGORY: I guess I just want to add that, from a reliability perspective, that it's still a pain when the lady's clearly in labor, but it was, somebody wrote latent phase and you can't get that to count. I mean, like could be five centimeters dilated and --

MS. MILTON: We changed it. We changed it.

CO-CHAIR GREGORY: Oh, thanks.

MS. MILTON: July 1st.

CO-CHAIR GREGORY: How is that changed exactly?

MS. MILTON: The inclusion area for labor. Latent. Yes. We heard from the field.

MEMBER FLANAGAN: I'm going to also add one more comment. It was included that there was a, what did they call it?

Interrelator reliability performed by

ORYX vendor. This is, I don't think it was recent. I think it was before the last version.

And the agreement rate was very high on the two areas that were of concern, which is active labor and gestational age. I don't know if you want to add anything on that.

MS. MILTON: Just that fact that it's from before. Yes.

CO-CHAIR GREGORY: Any other comments? Then I guess because there's been a change, we should vote on the, I'm sorry, I didn't see that, Jaleel.

MEMBER MAMBARAMBATH: I have the same thing as usual. Enrolled in clinical trials.

CO-CHAIR GREGORY: Enrolled in clinical trials.

MS. MILTON: It's gone.

CO-CHAIR GREGORY: So if, I think we should vote on this one because there's been a change. So --

MS. ROBINSON-ECTOR: Voting is open for reliability of measure 0469. It looks like

we're missing one vote. If everyone could resubmit their vote please. Great. All the votes are in and voting is now closed.

Ninety-six percent voted moderate, four percent voted low, zero voted insufficient. So for reliability of measure 0469, the measure passes.

CO-CHAIR GREGORY: Are there any new additions with regard to validity that you'd like to share?

MEMBER FLANAGAN: I think the, we already mentioned the change to the specs, the sampling and narrow performance interval.

I would like some explanation from the presenters on the, I don't know what page it would be, page seven I believe, on the distribution of outliers and what that was intended on showing.

I find it -- I've read it at least five times, and I went to the detail on page 27 and I still couldn't understand.

That 97 percent were considered

neutral, results not significantly different from target range, and 2.5 percent unfavorable. Can you explain that?

MS. MILTON: Stephen, are you on the phone? Ran the numbers.

PARTICIPANT: The operator says disconnected. He's not on the phone.

MEMBER FLANAGAN: I mean, my interpretation of this -- I'll keep speaking for a second. This looks oddly like an attempt to measure out performance and performance percentiles, and what I think this says with narrow variation is that even though you see this 10th, 25th, 50th and 70 and 50th percentile, then a lot of situations it's not statistically different from the target rate, and only in the very big outliers is it really statistically significant, which says that quartiles or percentiles may not be relevant for evaluating performance. But I'm not sure if that's correct.

MS. MILTON: As I look at this, this is what we submitted the last time. This is 2011

data because we weren't asked to change that. But I know we did some updated testing, so that's in a different section, I believe.

MEMBER FLANAGAN: Can you explain what this is trying to say?

MS. MILTON: I think based on what he did five years ago, that the majority were considered to be neutral.

In other words, there wasn't really anything that was distinct there and there was a very small number that were considered to be unfavorable or considered to be outliers. That most everybody fell within the range.

MEMBER FLANAGAN: If that's the case, then I would caution, I mean, as a caution to reporters like Leapfrog to not make distinctions that are not statistically significant in quartile reporting.

Not that that's the purview of this committee, but I mean, I think that's what you're trying to say here from the standpoint of validity.

Does this measure what it's supposed to, and is there meaningful difference between, you know, four different quartiles of performance?

MEMBER AUSTIN: Thanks for that feedback, Tracy. We can -- we'll consider that. Thank you.

CO-CHAIR GREGORY: So does the committee feel we need to vote on validity, or should we accept what's previously been accepted? I'm going to take that as accept what's previously accepted, and move to feasibility.

MEMBER FLANAGAN: I don't have much to say about this except that it is a measure that primarily, that does involve some degree of manual review.

However, almost all of the hospitals have been doing this for a while and I don't think there would be any change in burden. But I welcome any other comments on that.

MEMBER BAILIT: I would just say I think it actually decreases the burden with the

additional coding. I think the burden's gone down probably.

CO-CHAIR GREGORY: So I think we should vote on this. Yes ma'am?

MEMBER SHEA: I just have a question about the collection of gestational age that's used in this measure, and if you could just school me.

Is it at the time of delivery? Is it at the time of admission to the ICU? Is the gestational age on the birth certificate? What are the specifications around gestational age?

MS. MILTON: We like them to get it closest to the time of delivery because you could come in and be in the hospital for a few days and still be at 38 and 5 and eventually go to 39.

So we -- that's what we put in the notes for extraction, that it should, it should be at the time of delivery, and it's my understanding that's how the birth certificate data is collected.

It would be gestational age at the

time of delivery, and we do allow that as an allowable data source, if they're able to electronically retrieve that information from vital records reports from their states. Some states like California have that capability of doing that.

MEMBER SHEA: The reason why I ask is because I don't know all the specifics behind it, but I do understand that the way in which gestational age is now collected and reported by the CDC has changed over the last year, and that has, you know, then resulted in lower pre-term birth rates.

And so I'm just wondering, is this gestational age an estimate of gestational age based on, you know, EDC or is it by exam?

DR. MAIN: No. The, what the National Center for Health Statistics came out with, recommendation for all birth certificates in the United States is that it's based on a best obstetric estimate, which is generally ultrasound.

And ideally in agreement with last menstrual period. Previously they had used last menstrual period, which is not very accurate.

So this is a more accurate assessment of gestational age. And if you had the EDC, then it's easy to calculate the gestational age at the time of delivery.

MEMBER SHEA: So for the most part we might say is gestational age that's recorded on the delivery room record.

DR. MAIN: Yes. Yes.

MEMBER SHEA: Because if there's any opportunity for gaming here, that's where we see it. We see that the gestational age that's perhaps recorded on the delivery room record is not the same gestational age that's recorded by perhaps the healthcare provider in the notes, and we see a little bit of a difference coming through.

CO-CHAIR GREGORY: In general, most electronic records are automatically correcting. So if you come in at admission and you put 33 and

1, then on your delivery date, it automatically updates to 31, 33 and 2.

So it's pretty, they would have to blatantly go in and make a change for it to be, but I'm going to call for a question on feasibility.

MS. ROBINSON-ECTOR: Voting is now open for feasibility of measure 0469. All the votes are in and voting is now closed.

Sixty percent voted high, 40 percent voted moderate, zero voted low, and zero voted insufficient. So for feasibility of measure 0469, the measure passes.

CO-CHAIR GREGORY: We're going to do usability and use.

MEMBER FLANAGAN: So I think everybody knows that it's extensively used in quality check, hospital accreditation, hospital compare, hospital and patient quality reporting. It goes on.

And I think that what I found actually very revealing is Cindy's comment earlier about

how it has opened the door to understanding of maternity by a wider audience.

I actually thought it was a very profound statement because for so long, the area of maternity has not been really looked at. So I appreciate that comment.

The, as far as impact and improvement, it'll be very to see what happens when the additional numbers of hospitals come onboard.

Unintended consequences, I think we already talked about that. There's a section on expected findings, which I think we covered already, which was that there were some omissions in the original specs, lack of clarification on uterine surgery, which has now been clarified.

Those who didn't receive prenatal care, that where the gestational age was undetermined, that's been clarified, and the sampling issue. But I think it's, I think it passes on usability. That's my own opinion.

CO-CHAIR GREGORY: Comments? Vote.

MS. ROBINSON-ECTOR: Voting is now

open for usability and use of measure 0469.

Okay. All the votes are in and voting is now closed.

CO-CHAIR GREGORY: So now we need to vote on whether or not we would like to recommend that this be a measure --

MS. ROBINSON-ECTOR: Oh, just for the record, I need to read the --

CO-CHAIR GREGORY: Oh, I'm sorry.

MS. ROBINSON-ECTOR: It's okay. For usability and use, 84 percent voted high, 16 percent voted moderate, zero voted low, and zero voted insufficient. So the measure passes.

CO-CHAIR GREGORY: And now we would like to vote on the overall suitability for endorsement.

MS. ROBINSON-ECTOR: Voting is now open for overall suitability for continued endorsement of measure 0469. Option one is yes, option two is no. All the votes are in and voting is now closed.

One hundred percent voted yes, and

zero voted no. So for recommended continued endorsement of measure 0469, the measure passes.

DR. WINKLER: Yes, so the next measure to discuss, this is the eMeasure version of this measure. All right?

So if you look at the two measure information forms, they're identical when it comes to evidence and gap.

So if anybody, unless you have an objection, we can just carry over your prior discussion and votes and say, yes, they apply to this one too, which means we start out at the discussion of the eMeasure on the, you know, the specifications, what we know about reliability, validity, and the real, the things that are really specific to an eMeasure that haven't already been talked about in the other.

PARTICIPANT: Just a brief introduction to the eMeasure. It is the same complete description.

We are evaluating patients with elective vaginal deliveries or elective Cesarean

births greater than or equal to 37, or less than 39 weeks gestation.

This version of the measure was first specified in 2012, and it's updated annually. Those updates address the clinical updates to maintain alignment with the chart abstracted measures, so the evidence and the statements are exactly the same for the measure populations.

The updates also address updates to the standards for eMeasures, which are rapidly evolving. From a submissions standpoint in, as Reva was saying earlier, these specifications were developed in 2012.

Up until this year, hospitals have attested to CMS, that they're able to capture data on eMeasures, and they attest to their aggregate rates. The Joint Commission has not collected that data to date.

Beginning last year, we're accepting electronic submission of the measures, which is actually sending us raw patient data for calculating the measure rates.

And last year we had six hospitals, yes, seven hospitals submit PC-01, and in 2016, there are 69 hospitals that plan to send us data.

We'll receive that data in 2017, so we don't have it to look at yet. And we received the 2015 data this March. So we are still reviewing it on those seven hospitals that have submitted.

CO-CHAIR GREGORY: Would our discussants like to -- so we are accepting the evidence, and we are moving on to reliability.

Right, so we accepted the evidence, and we've accepted the gap, so we're moving on to reliability, and we're discussing this specifically because it's an eMeasure, and any comments the discussants would like to share.

MEMBER FLANAGAN: So I'm going to take reliability and validity almost together because, in this situation, I'm not exactly sure how to separate them, and I invite my statistical colleagues who are non-clinical to help me with this.

But my understanding in reading these two sections is that the way this gets validated, what you're trying to do is through a clinical upload, exactly replicate the measure without chart review, and with minimum burden.

I mean, that's the point of this. And so the question is whether what you do actually replicates what you think you're doing.

And what the presenter puts forward on this measure is a Bonnie testing of a simulated data set of 51 patients, and then kind of matching to see whether they pass or fail on the data included with each data element matched. This is the best of my ability of understanding. And then, and it in fact did pass. But I will just top there before I get myself into trouble of not understanding better than that.

CO-CHAIR GREGORY: Okay. Does anyone have any comments or questions? Yes, Matt.

MEMBER AUSTIN: Is the expectation that you would get the same rates? Because I noticed that the summary was that they match in

terms of passing or failing, but --

PARTICIPANT: The Bonnie testing tests the measure specification accuracy, is the exclusion excluding the codes we're looking to exclude.

It's not testing the performance rates on a representative population. So it's really just testing to make sure that the measure is working as expected.

We won't be able to compare measure rates until we have a significant data set with the eCQM to compare to chart abstracted. Does that answer your question?

MEMBER AUSTIN: Yes.

MEMBER FLANAGAN: I'm going to add a comment on this in that I can't imagine that the -- given that there's good, there's good correlation from element to element that this will in any way approach the variability and problems with reliability with sampling. And this takes away sampling, essentially.

PARTICIPANT: That's correct. There's

no sampling with the eCQM.

DR. WINKLER: I guess one question I have which sort of pulls all of it together is, now that you're going to be getting the data from the eCQM from 60 hospitals this year, seven last year, whatever it was, how are you, the Joint Commission, looking at that data for those particular healthcare organizations?

This is what they're giving you for their performance, and then everybody else is doing it the old way.

How are you looking at their results? Are they the -- are you just saying it's kind of the same as everybody else?

Are you doing, going to do any particular analyses to ask whether they are comparable results, or are we assuming they are, or are you going to keep them in their own bucket and look at them separately?

PARTICIPANT: I wish Stephen --

MR. SCHMALTZ: This is --

PARTICIPANT: Oh, good.

MR. SCHMALTZ: This is Steve Schmaltz from the Joint Commission. I think I can respond to that, at least a little bit.

For 2015, we actually received some data from hospitals for the PC-01 measure, and we're in the process of analyzing that right now. We didn't actually have hospitals that sent in both of them at the same time, so what we're doing is looking at data element by data element level.

Looking at data from the same hospitals for the previous year, and then looking at patterns of missing data or the type of data they tend to populate to kind of look at whether they're kind of reliable that way.

We actually have some stroke data that we're looking at where we can look at the same hospitals sending both at the same time, but we don't have that opportunity with the eMeasure or with the PC measure.

CO-CHAIR GREGORY: Kristi.

MEMBER NELSON: We've actually been

watching this both ways because we knew the electronic was coming and the caution is to, you have to educate your coders. That's where the biggest difference is.

CO-CHAIR GREGORY: Greg.

MEMBER GOYERT: Exactly. I think there's going to be a significant spread. It would seem to be valuable for the individual institution or symptom, or systems, rather, to look at both sets because, you know, after talking to the coders multiple times for every single fall, I never knew my parents weren't married.

But they informed me of that along the way. Because I see every single case that falls out in our system, and so there's going to be, if we don't get a bite at that apple to explain to the coder the minor error of their ways, there's going to be a disconnect.

CO-CHAIR GREGORY: Or the doctor. And --

MEMBER GOYERT: Coders.

CO-CHAIR GREGORY: And I guess they, with the eMeasure then, it should be a population measure and not a sample? Like, I mean, if it's an eMeasure, it's every single, 100 percent, cohort is the word I want. Okay.

MEMBER FLANAGAN: Let me add another comment on that. And if you're going to do side by side comparisons, it better not be a sample. It has to be the same denominator.

CO-CHAIR GREGORY: So I think if there are no further comments, we can, we'll do two back to back votes. One on the reliability, and one on the validity. Is that correct? Okay?

MS. ROBINSON-ECTOR: Voting is now open on reliability for measure 2829.

CO-CHAIR GREGORY: Do we need to vote again?

MS. ROBINSON-ECTOR: No. Okay. So all the votes are in and voting is now closed.

Eighty-eight percent voted moderate, eight percent voted low, four percent voted insufficient, so for reliability of measure 2829,

the measure passes.

CO-CHAIR GREGORY: The next one.

MS. ROBINSON-ECTOR: Voting is now open for validity of measure 2829. Option one is moderate, two is low, and three is insufficient.

DR. WINKLER: In terms of feasibility, I think this is one where the eMeasure is very different from the other, and the issues around eMeasures, and I think you've spoken to some of it.

So I do think this is one where they are different inherently, feasibility is frequently about the data source and the collection of the data.

MS. ROBINSON-ECTOR: All the votes are in and voting is now closed.

Ninety-two percent voted moderate, four percent voted low, and four percent voted insufficient. So for validity of measure 2829, the measure passes.

CO-CHAIR GREGORY: Okay. Discussants, can you please share with us information related

to feasibility of the eMeasure?

MEMBER FLANAGAN: Are folks in the room using this measure as an eMeasure? Kristi, you mentioned --

FEMALE 3: We're watching it, but I don't think we've submitted it that way.

MEMBER FLANAGAN: No, but you're actually doing it. I just emailed my data abstracter who for all of, we're not even going to intend to do it for two years, which is interesting because I actually thought, I love your beginning comments.

I've been told for years we're doing now what's called clinical upload, but that is not the same thing as eMeasure.

Clinical upload means that we create some -- we take as much from our electronic medical record, but we're still doing it according to the traditional way of submitting it, and there is no intention of going to anything, to using the eMeasure.

CO-CHAIR GREGORY: I guess I don't

know, as a committee member, I see that all of the eMeasures are, I mean, they're the goal, and they are corollary measures. But we don't know.

And so I think that if we've approved one measure and we've gone through the validity and the reliability of what we think the measure's going to do, I think the feasibility and usability has to be tabled until we have some more data.

DR. WINKLER: I think because it's part of the criteria for evaluation, you do have to make some comment on it.

CO-CHAIR GREGORY: Okay.

PARTICIPANT: You could say it's efficient.

DR. WINKLER: Possibly.

MEMBER FLANAGAN: Well, I heard our presenter say that there are six hospitals doing it. Can you comment on feasibility?

PARTICIPANT: I think I would be silly to say six hospitals is the country. I won't try and do that. You know, this has been used by

hospitals since 2012 through attestation, and so I don't have the numbers that the hospitals doing that.

I think the number seven for this measure is lower than we see for some other measure sets, but I think that's not necessarily related to the feasibility of this measure as it is the applicability of this measure to the population, or the vendor's interest in developing this measure for their hospital systems based on the number of available medical surgical measures.

So I think it's hard to comment on feasibility by saying six or seven is high or low, but there were seven hospitals that are able to submit this data to us.

We've also received feedback on the measure from EHR vendors and specifically coming to mind, Epic and McKesson and Cerna review the measures each year and provide their feedback on the measures, in addition to some other EHR vendors.

And one piece of advice we took from the vendors this year is that around capturing of gestational age and how many vendors are doing a calculation based on estimated due date.

And so not in the version of the measure we submitted, but in the next upcoming version we've included guidance on, that will, we would allow that calculation. So we are improving the feasibility of the measure over the time through feedback.

CO-CHAIR GREGORY: So maybe there should, I don't know. There should be a comment about that in the future. I'm going to -- unless there's any -- we've been tasked to vote, so if there are no other comments, I would like to vote.

MS. ROBINSON-ECTOR: Voting is now open for feasibility of measure 2829. It looks like all the votes are in and voting is now closed.

Eight percent voted high, 64 percent voted moderate, zero voted low, and 28 percent

voted insufficient. So for feasibility of measure 2829, the measure passes.

CO-CHAIR GREGORY: And then usability.

MEMBER FLANAGAN: I don't think there are any additional comments beyond the prior measure.

If we can get this right and believe it, which is really my shorthand way of saying that it correlates with the old measure; it should be very usable.

CO-CHAIR GREGORY: I'd like to suggest that we let the prior vote count. Can you let me do that? Okay. Okay, we're going to move to overall votes.

Whether we're going to, we are going to move to whether we would like to recommend this be endorsed for a measure. It's a one, two vote.

MS. ROBINSON-ECTOR: Voting is now open for overall recommendation of suitability for endorsement of measure 2829. Okay, all the votes are in and voting is now closed.

Eighty-eight percent voted yes and 12 percent voted no. So for recommendation of the overall suitability for endorsement of measure 2829, the measure passes.

CO-CHAIR GREGORY: So the good news is we're getting there.

We are now going to talk about measure 0480: Exclusive Breast Feeding and the subset measure Exclusive Breast Milk Feeding Considering Mother's Choice, and our developer is the Joint Commission.

So they get a two minute overview, and the discussants are Diana and myself, and we have no exclusions.

MS. MILTON: Good to go? Okay. For this measure, we're looking at single term newborns that have been discharged alive from the hospital, and of those newborns, we looked to see if they were only fed exclusively breast milk during the entire hospitalization.

This is a process measure. The goal is to improve the rate. The goal is not to get

to 100 percent, because we understand that there are going to be circumstances, maternal or neonatal, where you can't successfully breast feed.

And we do know there's going to be cases where mothers choose not to breast feed. But we've chosen to make this a very simply measure. It is your rate.

So we're not taking any exclusions out, and based on that, in our analysis we feel that an achievable goal should be about 70 percent of those newborns only fed breast milk during the hospitalization.

CO-CHAIR GREGORY: Okay. Diana.

MEMBER JOLLES: So as a maintenance measure, are we reviewing evidence? There's no change in evidence, and the evidence is rated -- is high quality. There's over 27,000 articles and systematic review. What?

CO-CHAIR GREGORY: And so we wanted to say that we will accept the prior evidence and talk about the gap.

MEMBER JOLLES: Okay. Moving on, unfortunately, there's still quite a performance gap with much room for improvement. This is felt by the committee and during previous discussions to be still rated as a high opportunity for improvement.

The goal of 70 percent is achievable. Less than 50 percent, the rate -- in less than 50 -- in over half of the Joint Commission hospitals that reported, they achieved less than 50 percent rates.

And then there were, in the 10th percentile, the hospitals were achieving 22 percent rate. Importantly, there are significant disparities on this measure with great room for improvement.

CO-CHAIR GREGORY: Right. And I'd just like to add that over time, they've gone from 166 hospitals to 1,400 hospitals reporting.

So as more hospitals report, the disparity or the opportunity for improvement is, increases. So if there are no objections, yes,

we have one. Yes, Jennifer.

MEMBER BAILIT: So I may be the only person in the America who's willing to say this publicly, but I am all for breast feeding. I get that.

But we are putting tremendous pressure on patients to breast feed when sometimes that's not appropriate. Who am I to tell someone who's working at a fast food job that she is a bad mother if she's not pumping in the bathroom on her only 15 minute break?

There are reasons that are good that people don't breast feed, and I think people get so caught up in this measure that they are not giving patients choice, even when the choice for that patient's life may actually be the right thing.

I'm not saying health-wise, but big picture, you know, she's got to keep a job. She's got to feed the family kind of picture stuff.

So I guess what bothers me here is we

don't have a balancing measure for this, and I don't even know what a balancing measure would be.

But I think this is a situation where the healthcare system puts tremendous moral, I will say, pressure because of this measure, and I think we just need to be aware of that when we talk about performance gap. That there's a price for that.

CO-CHAIR GREGORY: Cindy?

MEMBER PELLEGRINI: I was going to bring up two points. One is similar to what Jen said, and I don't know -- I know that Joint Commission, when they kind of put out these measures, I don't know how much, you know, guidance or resources they have for hospitals that are looking for how to balance some of these things.

And the other, I guess, question is, does the Joint Commission have data on hospitals that are -- have like Baby-Friendly status or other status, you know, sort of statuses and

whether the rates are variable by that type of program to show whether there's ways to improve.

I think hospitals struggle with programs with how to actually improve this metric in some way, and is there any sort of assistance with that.

DR. MAIN: There are some national toolkits on this. I think there's U.S. Breast Feeding is one that we've distributed to many of our hospitals, as well as the Baby-Friendly, which obviously is much more involved and expensive.

But U.S. Breast Feeding is a very good source, very specific about the steps. And it has some sample language, because I think this is all about language, as Jen was saying, about how you present it to your staff, as well as to women.

You know, there are certainly hospitals with large minority groups that do well over 70 percent, so this is not just a minority versus, you know, more privileged population

issue.

But I think it is about how our line staff understand and present, and go about their daily work with it. I think that's an opportunity.

MEMBER PELLEGRINI: Can I, sorry, can I bring up one other point? Just because, when Jen, you just mentioned, what would a balancing measure be.

One of the things that we noticed, because we just went through this, the Baby-Friendly process, and we noticed actually that our infant fall rate went up in that process. Infant falls.

And you know, part of Baby-Friendly is a little more intense in terms of the rooming in and the amount of time the babies can spend in the nursery or not really having a nursery.

And there's a lot of implications that are sometimes difficult for hospitals, and we noticed that our fall rate increased, and I don't know in the future if that's something that is a

consideration for a balancing measure.

CO-CHAIR GREGORY: Okay. Thank you. I think it's time to vote on whether there is an opportunity for improvement.

MS. ROBINSON-ECTOR: Okay. Voting is now open for performance gap of measure 0480.

CO-CHAIR GREGORY: Are you good?

MS. ROBINSON-ECTOR: Oh, I'm still waiting for more votes to come in.

CO-CHAIR GREGORY: More votes. Vote again everybody.

MS. ROBINSON-ECTOR: Great. Here we go. Okay. All the votes are in and voting is now closed.

So 75 percent voted high, 21 percent voted moderate, four percent voted low, and zero voted insufficient. So for performance gap of measure 0480, the measure passes.

CO-CHAIR GREGORY: Okay, we're going to talk about reliability. And is there anything new?

So it was updated to ICD-10, and

they're removing the sub-measure of exclusion of mothers who declined to breast feed, and this was actually based on stakeholders asking them because they felt that this was in -- too much burden to get this extra data. Do you want to add anything?

MEMBER JOLLES: I would just add that the reliability testing came out at 97.53 percent, which would rate it at a moderate.

CO-CHAIR GREGORY: So should we vote? Raj?

MEMBER WADHAWAN: Does that include patients who are in the neonatal ICUs as well? Term infants who are moved to neonatal ICUs, or is it just newborns?

MS. MILTON: Yes. Any newborn admitted to a NICU would be excluded from the measure, as well as those with galactosemia or receiving TPI or if they expired, or if they had a length stay greater than 120 days.

CO-CHAIR GREGORY: Although there is data that it's better for the NICU babies to get

breast milk, so I just wonder why we would --

MS. MILTON: We totally agree, but we're only focusing on the healthy term newborns, the singletons.

CO-CHAIR GREGORY: Okay. All right. Kristi?

CO-CHAIR GREGORY: What about the special care nursery?

MS. MILTON: Unless it meets the AAP definition of the highest level of care, that would be a no.

So we have to have specially trained staff, all of the appropriate equipment to take care of complex and extremely ill newborns in order to be considered a NICU.

MEMBER NELSON: I thought the measure was something about life saving treatment.

MS. MILTON: We have adhered to the AAP definition of the highest level of NICU care.

MEMBER NELSON: Okay.

CO-CHAIR GREGORY: So shall we vote on reliability?

MS. ROBINSON-ECTOR: Voting is now open for reliability of measure 0480. Okay. Looks like we're missing just one vote. If everyone could resubmit their vote please.

CO-CHAIR GREGORY: Two people left.

MS. ROBINSON-ECTOR: That's included. Great. Thanks. Okay. All the votes are in, and voting is now closed.

CO-CHAIR GREGORY: Okay, validity. Oops, I'm sorry.

MS. ROBINSON-ECTOR: It's okay. So 91 percent voted moderate, nine percent voted low, and zero voted insufficient. So for reliability of measure 0480, the measure passes.

MEMBER JOLLES: So for validity, this is rated at the moderate, I believe, because it's being evaluated at the data element level, of note perhaps to validity, when we look -- when the data is presented on the exclusions, it's of interest that there were zero galactosemia and zero pre-term exclusions, which, but otherwise --

DR. WINKLER: The only thing I would

notice is they did add additional empirical testing of the measure score. So that does put the possibility of a high rating on the table.

MEMBER JOLLES: Oh, okay.

CO-CHAIR GREGORY: So, shall we vote? Great.

MEMBER JOLLES: Any discussion on that?

CO-CHAIR GREGORY: There's a couple up. Okay. All right everyone. Greg.

MEMBER GOYERT: Just want to make the note that I'm going to present alone. When you look at the validity testing, the phrase is, the measure score correctly reflects the quality of care provided.

It does not. It reflects what patients choose to do. It doesn't reflect the counseling, the education, the time investment, so on and so forth.

So this is a measure where the institution is being evaluated on the basis of the choices that our patients make. It's not

necessarily a reflection of the quality of the care that we provide. Sermon's over.

CO-CHAIR GREGORY: Okay. Cindy, was your --

MEMBER PELLEGRINI: Yes, can I just, I just want to note that I think sometimes we hold this measure to a higher standard than everything else.

Almost everything we do with patients reflects their choices, whether they decide to have a procedure, whether they decide to take their medicine, whether they decide to come to the doctor at all. So I'd encourage us to think about that fact.

CO-CHAIR GREGORY: Sindhu, is your back, your card up? Sheila.

MEMBER OWENS-COLLINS: I just want to echo what everybody else is saying. And I mean, there's a very strong cultural and ethnic component to this that I'm not sure is included, that I think we should be sensitive to and address it and not penalize facilities for that.

CO-CHAIR GREGORY: Tracy?

MEMBER FLANAGAN: Right around the time about a year or two ago, I think it was Joint Commission was proposing an excellence standard according to performance on this measure.

And when I looked at what -- that there were hospitals in the 100 percent, I don't believe that a hospital can be at 100 percent, and so it made me question what kind of oversight Joint Commission was doing on this.

I can tell you that in our hospital system, the way we submit this is every feeding from every nurse, the whole hospital.

So I had lots of doubts about the validity of the people who were up beyond 95 percent. So I'd love to hear some comments on that.

MS. MILTON: I don't believe there are that many that are above 95 percent. There was one hospital out in California, and we interviewed them, but they're about the only one

that I'm aware of.

Again, we're not looking at 100 percent. I want to stress that. Seventy percent is where we feel it should be, where a hospital should strive to achieve, because there are personal preferences, and there's going to be conditions.

Mother's HIV positive, it's not a very common thing. When we did an analysis, a 12 month analysis, only two percent of the exclusions were due to medical conditions.

So we're not looking for anyone to get anywhere near 95 percent. But I know that there are a lot of hospitals that are above, not a lot, but there are some that are already above 70 percent.

MEMBER FLANAGAN: Let me restate that. I think that for those hospitals that are super high, I mean, maybe we choose 90 percent, maybe we choose 85 percent, I think it would be interesting for Joint Commission to perhaps prioritize them for audit.

CO-CHAIR GREGORY: Juliet. Oh, you wanted to say something?

DR. OWENS: We actually do audit them. We have, I alluded to this earlier, but we have actually a pretty rigorous -- not pretty, a very rigorous process for checking the quality of the data that we receive.

Healthcare organizations collect the data. They submit them to like a, we call them performance measurement systems, but basically they're vendors.

They run all of these quality checks. They are contracted with us. They are contracted with the hospital.

They're required to do inter-rater reliability and that kind of a thing. We're pretty confident that the numbers that we're getting are good. And we do check them on a routine basis.

CO-CHAIR GREGORY: Juliet.

MEMBER NEVINS: I'll start by asking a question, because I don't know the answer in

terms of the percentage of women who start breast feeding in the hospital who then continue for even a week, you know, two weeks past that point.

And the reason I'm asking this question is because I'm sort of piggybacking on Jennifer's comment. I think it's something that as a society we really need to strive for, because it's so important, and the health benefits have been so well demonstrated.

But patients don't want to start something that they can't finish. And we need, we, you know, America, we send our patients back six weeks after vaginal delivery, and eight weeks after a C-section.

Most of the time they can't, they can't breast feed in that work environment, even with the new laws requiring that a place for them to pump and store be provided.

So you know, in terms of holding hospitals accountable for breast feeding, even though this measure is only for measuring members or patients who breast feed in the hospital, it

sort of has -- it extrapolates beyond that time to moms who breast feed for at least six weeks, I think. So I just wanted to throw that out there.

CO-CHAIR GREGORY: But the data is clear that six weeks is better than none, so I would -- but I do appreciate your perspective. Naomi? Diana.

MEMBER SCHAPIRO: So when I was in nursing school, which was like 1972, there was only one woman in my entire OB rotation who breast fed.

So we, you know, we've done a really like an incredible job since then, and I've worked all my entire, pretty much present career with very low income women and families, and I would say that the more this becomes standard, the more conditions change to enable people.

That there are a lot of women in not very high status jobs who actually can pump now. And every woman goes home from the hospital with a pump if she's nursing, breast feeding. A really good pump.

You know, so I feel like if we relax the standards in some ways, then we relax the conditions that allow women who are in more difficult circumstances to absolutely choose to breast feed for longer, and I agree they need that.

So I, that's kind of, I understand choice, but choice is really dictated by circumstances, and so I really feel that this measure helps keep the circumstances there that allow more women to breast feed if they want to for longer. And I really support it.

MEMBER MOORE: I'm going to build off two comments that were made. Essentially they made my comment, but I wanted to ask additional pieces there.

The concept of measuring exclusive breast feeding, or the intentions of exclusive breast feeding, really doesn't apply to long term breast feeding, and I think that this measure is what, we assume that that's what's happening, and I think that we need to look at how we actually

measure breast feeding long term, because that's where the true benefits are, not that you have intentions during that discharge out of the hospital. The other piece, I love your comment Naomi, but I'm really struck by Jen's comment, too.

And I feel like there's this, there's this balancing between the two. And knowing that I'm in the Medicaid space and knowing that 50 percent of births on Medicaid, they're low income women, they're going back to work way before four, six weeks, and they're going into jobs where being able to pump, those accommodations are not being made for them.

And this notion that we put all this pressure on women that they're not good moms, that they're not getting what's best for their newborns, coupled with this notion of, well, let's keep that threshold at 70 percent to help move the nation forward.

I mean, I'm really challenged by the two comments, and it's hard with this measure,

and to expect that we'll be at 70 without considering the patient's situation as occurring right now.

CO-CHAIR GREGORY: Okay. So I see one, two, three, four cards up. It's 10 to 2:00. The meeting is supposed to adjourn at 2:00 and we still have one more measure. So keeping that in mind, Diana. Diana you're next.

MEMBER JOLLES: Well, I just, I'm only going to say this because I think it's important in general as we debrief over our entire portfolio.

So this measure is one of two measures that affects population health. We're looking at a large population of over four million child bearing women.

The minority of our measures are addressing those, that population movement. This measure is affecting life course health, chronic disease. This issue of women being pressured is a lack of process.

So I just want to raise awareness

about this. The Baby-Friendly hospitals have a very good process, and it is easy for people to opt out and sign and never be harassed again.

So if people are being harassed, it's because they're in systems that aren't embracing improvement and process. I must sit as a midwife and say that I ran a service in Washington, D.C. that served African American women and we had the highest breast feeding rate in this entire city across all socio-demographic groups.

We had group prenatal care, and most importantly, peer counselors. And we did not pressure anybody. And so this is a process measure that has extreme supply-sensitive provider preferencing, and poor quality.

So please ride this measure out and watch it until we can retire it and then get into those issues.

CO-CHAIR GREGORY: Okay, Deb.

MEMBER KILDAY: Okay. Apparently I need a nap. I'm just going to sort of echo on what you were stating.

Having the observation to look at how hospitals perform and how they use their quality initiatives, a lot of hospitals are talking about breast feeding.

I do hear you about patient preference, and maybe it's the patients, but from my clinical experience and from my observational experience in consulting with all these hospitals, I see a tremendous clinician culture, nursing predominantly, including pediatrician and obstetrician education and knowledge, and some resistance between those three sort of effects in furthering our process.

So I see less with patients, but I see a lot of babies in nurseries these days, which we don't think we're supposed to see.

I see a lot of formula sitting out all over the place, which we're really not supposed to see. And then when I do chart abstraction quite frequently when I do these quality assessments because, you know, you really get into the weeds by opening the chart.

And we're feeding babies a lot of things we shouldn't be feeding them, and when you ask the parents, because I do that on my assessments, they're like the nurse told me that it would probably be better if I get some rest.

So I'm going to echo your statement that we have tremendous process problems within our culture of healthcare, and I'd like to continue to take a look at that.

CO-CHAIR GREGORY: All right. We have two more comments. I hope they're pertinent to validity. Cindy.

MEMBER PELLEGRINI: This is a challenging issue, but that is not the measure's fault. We can do better. We need to do better, and the measure can help track our progress to getting there.

But when we have patients who smoke, we don't say, well, they're addicted to tobacco and we should change the measure. We say, do we have the supports, do we have programs in place, do they have access. We don't say -- we don't

throw our hands in the air.

CO-CHAIR GREGORY: Okay. I think this was all very exciting discussion, and we will not vote on whether this is --- if the measure has validity.

MS. ROBINSON-ECTOR: Voting for validity of measure 0480 is now open. It looks like we're looking for 23 votes, so we're missing one vote, if everyone could resubmit.

CO-CHAIR GREGORY: Okay.

MS. ROBINSON-ECTOR: So 22 votes are in, so voting is now closed.

CO-CHAIR GREGORY: Okay.

MS. ROBINSON-ECTOR: So 36 percent voted high, 55 percent voted moderate, 9 percent voted low, and 0 voted insufficient. So for validity of measure 0480, the measure passes.

CO-CHAIR GREGORY: Okay. With regard to feasibility.

MEMBER JOLLES: Feasibility is rated high. It's used by Quality Check, Joint Commission, the Hospital Inpatient Reporting

Program, eClinical measure, CMS9, and meaningful use. So that would be a high.

CO-CHAIR GREGORY: I would -- I have nothing to add to that, so we can vote on feasibility.

MS. ROBINSON-ECTOR: Voting for feasibility, measure 0480 is now open.

CO-CHAIR GREGORY: We're going to vote on usability. Any comments on usability? I'm sorry. Keep me on it.

MS. ROBINSON-ECTOR: For usability --

CO-CHAIR GREGORY: Oh, sorry.

MS. ROBINSON-ECTOR: Okay. So all the votes are in and voting is now closed. Seventy-eight percent voted high, 22 percent voted moderated, zero voted low, and zero voted insufficient. So for feasibility of measure 0480, the measure passes.

CO-CHAIR GREGORY: So with regard to usability, it is currently being used, it is publicly reported, and it is part of accountability programs. Would you like to add

anything?

MEMBER JOLLES: Currently until the eMeasure is adopted, there is lack of public performance data available to consumers, and I would add that an important part of usability is improvement, progress, and the ability to have impact, and as we already spoke about, this is a high impact measure.

CO-CHAIR GREGORY: Any comments? Okay. We shall vote on usability.

MS. ROBINSON-ECTOR: Voting is now open for usability and use of measure 0480.

CO-CHAIR GREGORY: I'm sorry. Ashley?

MEMBER BELL: Just quickly, it's not probably going to affect this, but I didn't know when to bring it up and maybe in more in gaps actually, but I know this measure was discussed with the Medicaid MAP.

It was last, I guess, yes, last year. And it was really debated, like we knew that breast feeding was important, but I think that exclusivity and maybe Carol, if you remember, you

were on the MAP as well.

I think that's what was problematic about it, that there was concern with kind of current practice I guess, maybe with pediatricians wanting to intervene with weight loss in the newborn, kind of pushing the supplementation that exclusivity was really impossible to achieve.

I know that the data are about 50 percent, but if we just think about how we can improve this measure given those kinds of concerns or think about adding another measure that really just captures initiation versus exclusivity.

CO-CHAIR SAKALA: So what I recall is there was nobody in the room who was providing care to childbearing women, and the concern was a focus on the coercion aspect, and a lack of recognition of the system issues and the potential to have very good process, and also that the -- it was added as a candidate measure the day of, rather than included in the list.

So I think it's going to be included in the list later on this month, and let's see what happens at that point in time.

It -- I was a little concerned about the way the discussion played out because of lack of understanding of these -- an impression without having firsthand experience.

CO-CHAIR GREGORY: Okay. So we're going to vote on usability and use.

MS. ROBINSON-ECTOR: Voting is still open for usability and use of measure 0480. All the votes are in and voting is now closed.

Sixty-one percent voted high, 35 percent voted moderate, 4 percent voted low, and 0 voted insufficient information. So for usability and use of measure 0480, the measure passes.

CO-CHAIR GREGORY: And now we'll vote on whether or not we would like to put it forth for continued endorsement.

MS. ROBINSON-ECTOR: Voting is now open for overall recommended suitability for

continued endorsement of measure 0480. Okay. All the votes are in and voting is now closed.

CO-CHAIR GREGORY: All right, we're going to count and make sure we still have a quorum.

MS. ROBINSON-ECTOR: So 91 percent voted yes, and nine percent voted no. So for recommended suitability for continued endorsement of measure 0480, the measure passes.

CO-CHAIR GREGORY: So just a second. Are we good with quorum? We're okay. So I'm going to ask Reva to help me through this because it is a corollary measure, and I will try to only do the votes we need to do, and have the discussion we need to have.

DR. WINKLER: Okay, so let me help drive this one. Okay. Evidence and gap, it's the same information, so we can carry it over from the one you just discussed. We're done. Okay.

Reliability and validity really are as we discussed with the prior eMeasure, the

specifications that are unique to the eMeasure, as well as the Bonnie testing.

So we do want to have a conversation about that. The feasibility is probably going to be the same as it was for the other eMeasure, right?

CO-CHAIR GREGORY: Right.

DR. WINKLER: If you like, we can carry that one over. Usability you felt was still, you know, a little hard to get your hands around. Probably going to be the same as the other. We could carry that over.

So if you would just talk about, you know, the specs for this measure, the Bonnie testing, and then we'll do an overall, that might get us there quickly.

CO-CHAIR GREGORY: Can you help Diana out?

MEMBER JOLLES: So beginning with quality construct?

CO-CHAIR GREGORY: The reliability and the validity for the eMeasure.

DR. WINKLER: Really the specifications and then the Bonnie testing.

MEMBER JOLLES: Okay. So this is, this is easy to get through. Its HQMF specifications have been made. They've been vetted through USAC.

Value sets exist, Bonnie testing occurred, which was demonstrated to be reliable. And then it is a legacy eMeasure that's being used for meaningful use. It's already got SNOMED mapping. So the quality construct is there, is present.

CO-CHAIR GREGORY: Okay. So we will vote on the reliability. It's a one, two, three, right?

MS. ROBINSON-ECTOR: Yes, it's moderate or --

CO-CHAIR GREGORY: So it's moderate, low, or insufficient. Moderate is one, low is two, and insufficient is three.

MS. ROBINSON-ECTOR: Yes. So voting is now open for reliability measure 2830, or

2830. Option one is moderate, two is low, and three is insufficient.

CO-CHAIR GREGORY: You're looking for 20 votes I believe.

MS. ROBINSON-ECTOR: Okay. So all the votes are in and voting is now closed.

CO-CHAIR GREGORY: Fabulous, and --

MS. ROBINSON-ECTOR: 89 percent voted moderate, 5 percent voted low, and 5 percent voted insufficient, for reliability of measure 2830, the measure passes.

CO-CHAIR GREGORY: Would you like to comment on validity?

MEMBER JOLLES: Sure. Bonnie testing was conducted of the eMeasure, 528 cases passed at 100 percent.

CO-CHAIR GREGORY: So we'll now vote on validity.

MS. ROBINSON-ECTOR: Voting for validity of measure 2830 is now open. Like we have, okay, all of the votes are in and voting is now closed.

Seventy-nine percent voted moderate, 21 percent voted low, zero voted insufficient. So for validity of measure 2830, the measure passes.

CO-CHAIR GREGORY: So our last vote of the day is whether or not this will be recommended for consideration for endorsement, and it's a yes or no vote. You said we're carrying it over? That's why I had --

MS. ROBINSON-ECTOR: Voting is now open for overall recommended, recommendation for suitability of endorsement for measure 2830. All the votes are in and voting is now closed.

CO-CHAIR GREGORY: I would like to thank everyone and commend you for your --

DR. WINKLER: And, yes. I'll take over for this one. No, again, I echo Kim, and I can't thank Kim and Carol enough for what they've done. There were a couple questions about, okay, now you've done this, now what.

What you would've done is act as a proxy to the multi-stakeholder membership of NQF

and the public at large.

And so you have made the sort of first pass recommendations back to NQF about which measures should be endorsed.

So this is the first step, or an early step, okay, going forward. So we will be writing a draft report, and then we will, with your recommendations, and we will be soliciting public comment. Okay?

MEMBER OWENS-COLLINS: When will that be?

DR. WINKLER: We'll announce it to you. It'll be in about a month. It's in the month of June pretty much.

And so then anyone is welcome to comment and submit written comments, then we will collate those comments and come back to you.

We have a scheduled meeting at some point, a call at some point. There's a date already set. And we'll discuss and ask you to respond to those comments and perhaps if they make any influence in your decisions of making

your recommendations.

After that, those recommendations, after considering the comments, will go for an NQF member voting, and that's in August. And then it goes to our Consensus Standards Approval Committee, which is a subcommittee of the Board of Directors, who looks at the more detailed aspects of the process and leading up to granting the final endorsement by CSAC, by recommendation of the Board by ratifying it.

So through the summer, we'll be sort of finishing the details. So you're at the beginning of this process that'll go over the next couple of steps. So I know somebody had a few questions. So I wanted to go over that.

MEMBER WADHAWAN: What is the process after reviewing public comments? If there is a desire to re-look at the measure and change the recommendations. Is there a repeat vote?

DR. WINKLER: Yes.

MEMBER WADHAWAN: Is it by phone or --

DR. WINKLER: Yes. You would be

repeating the vote, discussion and vote. Yes. Absolutely. Okay. So I just wanted to respond to that before we lost too many more people.

Now I understand that we do have someone on the phone who wants to make a comment. Or in person, fine. And for our last opportunity for public comment today.

MS. SANTA-DONATO: I'll be brief, I promise. Is this on? Hello? Okay. Hi, I'm Anne Santa-Donato, and I represent the Association of Women's Health Obstetric and Neonatal Nurses.

I'm the director of Obstetric Programs, and on behalf of the organization, I just want to thank you for the opportunity to attend this meeting and to let you know that our organization fully supports the mission and the work of the NQF as well as the measures that were endorsed today.

I just want to provide you with a very brief update about AWHONN's Women's Health and Perinatal Nursing Care quality measures.

We've developed a series of 12 measures over the last few years, and we are now in the testing phase, and two of those measures are currently being tested.

One is immediate skin-to-skin care following birth, and the second one is the continuation of, duration of continuation of interrupted skin-to-skin care during the birth hospitalization.

And those measures are being tested through the NPIC hospitals, particularly some of the CWISH hospitals that are part of NPIC. In addition to that, we have a maternal fetal triage index, which is designed to help the, to facilitate the very first assessment of women who are coming into labor to be evaluated, modeled after the Emergency Nurse's Association scheme for initial assessment of the patient to prioritize care.

In order to be able to test that measure, a tool was necessary to be developed that had some standardization to it.

So that tool was developed and tested, and so that measure will be ready for testing, hopefully within the next few months.

So I just wanted to provide you with those updates, and thank you so much.

DR. WINKLER: Okay. At this point, what we always like to do before we finish up is kind of look back over our portfolio to kind of see where we are.

You know, we've talked about several new measures. Not all of them were recommended by you, but I think in the course of looking at the various measures, the idea that there are probably areas that aren't being measured that represent quality problems within that portfolio.

And you have the opportunity to make some recommendations on those areas. We call them gaps. But one way to think about them is really around where are the quality problems that would be amendable and responsive to measurement.

Not everything can be fixed with a measure. But our focus is measurement. And so

as we look at our portfolio of measures say in reproductive health and antenatal care, it seems we're a little light on antenatal care measures.

And hello. Go back. Thank you.

And again, I would, you know, with three new contraceptive care measures, it starts to really add some weight to the reproductive health area.

Realizing that there's a significant overlap between whether we categorize something as reproductive health versus women's health, we do have measures in general women's healthcare, and we only pulled the ones that have very specific focus on reproductive health.

But your thoughts around the measures we have and the measures we don't have. And perhaps what might be happening out there in your world of the things that you're looking at and attending to and finding that they are problems that you're trying to address through, perhaps local measurement or something like that. Thoughts from anybody? Go ahead, Cynthia.

MEMBER PELLEGRINI: A quick plug for preconception care, which is of course a subset of women's health, and March of Dimes is starting to have some conversations with CMS about taking some of their existing preventive measures and pulling them together into either a formal or an informal preconception measure set.

The interesting thing about what else would be missing from that.

MEMBER FLANAGAN: Just a couple comments along the same lines. I think an area that really needs some focus is perinatal depression.

And the second area that I think we could do better in and actually could even create a measure or entertain a measure is identification of intimate partner violence or DV.

CO-CHAIR SAKALA: Yes. So I have had some conversations with our reproductive health team, and this is a little bit of an echo of what Lorrie Gavin said they're working on that I just

would like to put a support for it.

Patient-reported experience of contraceptive composite measure, whether women felt respected, whether they were informed and whether they experienced shared decision making, which would be a nice balancing measure for those.

And also in more broader settings than just once at our, extensively focusing on contraception, a measure to track whether women were screened for pregnancy intention and desire to use a contraceptive measure.

MEMBER AUSTIN: Yes, and I may be sort of stating the obvious, but looking at the pregnancy measures, I mean, those two both ran into some challenges yesterday.

So there seems to be a lot of good discussion around what would be a meaningful prenatal care measure or the elements of that. So that might be worth exploring.

MEMBER BAILIT: So having publicly come out against breast feeding measure that

exists, I'd like to see something better.

You know, is it a pediatric measure, some breast milk at six weeks, at six months, and our pick of time. And it should be some and not exclusively would be my thought.

MEMBER KEATS: So is depression screening in the perinatal period being worked on, did you say?

Or, I mean, that's not a measure that's in development right now. Or Tracy, you're the one that brought that up.

MEMBER FLANAGAN: The U.S. Preventative Services just came out with recommendations in January endorsing this with, I think, moderate strength evidence.

MEMBER KEATS: Yes, it was --

MEMBER FLANAGAN: But as far as I know, I don't know of any measure developer at this point, while there's a lot of interest.

MEMBER KEATS: Yes, okay. So a developer has not been identified is what you're saying. Okay, great. Thank you.

CO-CHAIR SAKALA: I just want to say that the PCPI set has a composite postpartum measure which does include depression screening.

And I'll put in a general plug for the AWHONN's that Anne just discussed, and the PCPI set that have a lot of great potential for clinician level measures that are, for the most part, languishing for lack of testing support.

DR. WINKLER: Who does?

CO-CHAIR SAKALA: AMA-PCPI and AWHONN for nursing measures.

TL: What is PCPI? I've not ever heard of that.

CO-CHAIR SAKALA: Physician Consortium for Performance Improvement. It's with the AMA, and in this particular case, it was the collaboration with ACOG and NCQA in a multi-stakeholder process that worked those measures out.

DR. WINKLER: Okay. Cindy, are you up now?

MEMBER PELLEGRINI: I was waiting to

see if somebody else would get it first. Opioids prescribing during pregnancy, screening for pregnancy when prescribing, all that.

DR. WINKLER: Yes, right. Sindhu.

MEMBER SRINIVAS: I was going to say that too, and then I also, looks like we're on the same page. Now I know, we have a similar name so it goes along.

The other thing I was going to say is sort of thinking about, and this is a more difficult one I think, but thinking about all the counseling that's supposed to happen during prenatal care, and how it's not really -- it's not easy for clinicians or providers to do that counseling in terms of like nutrition counseling, weight gain, contraception, all the stuff that's supposed to happen sort of during the course of prenatal care, and somehow coming up with a measure that would be -- allow us to kind of really push that movement forward.

MEMBER FLANAGAN: One of my other committees that I sit on is the International

Committee for Healthcare Outcomes, which is the patient centered outcomes around maternity.

And the final set is just about being finalized right now, and they're looking for testers of it, and it does include women's perspective of their own prenatal care, which I actually think is probably a better indicator of good care than, you know, some of the more, what we might call medical ones.

DR. WINKLER: Tracy, who was doing that?

MEMBER FLANAGAN: International Consortium of Healthcare Outcomes. Yes. The last two years has been dealing with pregnancy specifically.

DR. WINKLER: Hold on. Let me, let me get it specifically. Because I don't want to get it wrong. Let me get it on my iPhone.

MEMBER SRINIVAS: Can I ask another question?

DR. WINKLER: Sure.

MEMBER SRINIVAS: The PCPI that you

just mentioned, is that, what, like once that's developed, like what happens?

DR. WINKLER: Well, it, this is really fascinating because it just shows you some of the silos we all live in.

The PCPI actually has been responsible for the development of both the measures used in all of the physician, you know, PQRS. Yes.

That's where they all came from, and the fact that obstetrics is sort of, doesn't play in that field is interesting.

And so they have developed with a lot of the specialty societies, clinician level measures. I don't know what the status is because the PCPI measures are kind of languished.

They were looking for opportunities to test them, and were not able to, and it's been a while since I've heard what their status is. But they did create a set of, I don't know if it was 5 or 6.

CO-CHAIR SAKALA: 10.

DR. WINKLER: Oh, was it 10? Okay.

Measures. I haven't seen them in a while. We'll see if we can figure out what may or may not be happening. But again, the challenges are testing these measures.

As you've seen, this is not a minor undertaking. And it tends to be the major challenge for any measure development, is getting the adequate testing and the resources and the people involved. Ashley.

MEMBER HIRAI: Comment on that. So that's actually -- it's the Behavioral Health Risk Assessment, and the steward is the PCPI in addition to NCQA and ACOG, and it's actually a prenatal, it's not a postpartum measure.

I think something parallel to that could be developed in the postpartum period. It does capture depression screening, alcohol use, tobacco use, drug use, which would include opiates, and intimate partner violence screening.

So it is this composite measure. It's been part of the Child Core Set for Medicaid. I don't know how many years. I think it's newly

added.

So only two states reported it in the last MAP report, so that was fiscal year, reflecting fiscal year '13 data.

So there's some problems reporting it, but I think everyone was really interested in that, and it captures a lot of different domains that then, and that may reduce some of the measurement burden versus parsing these out as separate measures.

But then, what are you truly reflecting, and what do you have to improve? Is it one of those, or is it a couple of them? So there's some problems, I think, with the composite approach.

But yes, if anything we can do to encourage them to continue seeking that NQF endorsement, and to extend that to the postpartum period, because a lot of those same screenings are relevant for the postpartum population as well.

MS. ALLEN: So the Behavioral Health

Risk Assessment is in the Child Core Set. However, states have been having some difficulties reporting that measure because it relies heavily on chart data.

Over the past several years, when it became -- it was included in the Core Set in 2013, and so only two states reported it. For 2014, only four states had reported the measure.

So CMS is really trying to increase data on the measure, but they're having some challenges with that.

MEMBER HIRAI: Thanks, Nadine. And just to move us along maybe, in Kim's spirit, I'm just going to say one thing about the perinatal measures then, and actually Cindy mentioned this yesterday with the concern about postpartum not continuing it or retiring it without having a replacement.

The same can be said for risk appropriate perinatal care. I thought that was kind of a revelation that we didn't pass that new measure proposed, and then Elliot is, didn't

apply to continue the risk appropriate perinatal care, so now we have no NQF-endorsed measure capturing that, and it's a very important measure, I know to March of Dimes.

And we know that it causes death when babies are not born in the appropriate facility. And I guess I just think that's going to be a gap now that we don't have a measure for that.

I will just encourage colleagues at CDC who have taken some ownership of that to maybe start to work on an application.

CO-CHAIR GREGORY: I guess I would just like to make a plug for outpatient measures of the content of care, as well as overutilization.

I hesitate to say that as an MFM, but I definitely think that there's some overutilization going on. And a measure of maternal morbidity, or a measure of total outcome of care.

DR. WINKLER: Do you know if anybody is working on something like that? Has anybody

even thought about what might be included in that?

MEMBER HIRAI: I think actually CDC is working maybe on an application for severe maternal morbidity using claims data.

DR. WINKLER: Okay. Interesting. Okay.

MEMBER HIRAI: Or hospital discharge.

CO-CHAIR SAKALA: Population level or health plan level. It doesn't go down too low, right?

MEMBER HIRAI: Yes, it can go to facility. Yes.

MEMBER BAILIT: NICHD with the APEX trial, did put together maternal morbidity outcome measures that are risk-adjusted.

They are based on chart review though and not on diagnosis codes, and for a variety of complicated reasons, it's hard to get the NICHD to be a sponsor.

So to the extent that somebody would want to work as a partner and be a developer, I

think those have a lot of potential.

DR. WINKLER: When we report your recommendations out, it's in a report that will contain some of these other things acknowledging the issues in the portfolio. Sheila, you wanted to say something?

MEMBER OWENS-COLLINS: All right. Couple of things. Okay. So I agree that the postpartum care exam is something that we should look at, but I just have tremendous angst with the way it's being reported now, and the consequence, the financial consequences that it has had on health plans because of the narrow window.

So I would strongly encourage that we really take a look at that, widen the window or make it a little bit easier to obtain.

Because it is used as a carrot or a stick for the health plans. Also, I'm involved with a grant from the state of Maryland to look at gestational diabetes in women in the follow-up care, and in this process, I have found that

there is a gap, and there are lots of opportunities.

Even for the six week exam, six week postpartum exam to look at those women because of the high risk of developing Type 2, the issues of coordination of care between the OB and the PCP and the transition of care.

So I would recommend that we take a look at that. And lastly, going back to that postpartum, you know, and if we could be more specific in terms of if we're going to stick to the narrow window of specific conditions that are amenable to looking at specifically at six weeks.

And the gestational diabetes one is a piece. The neonatal admission rate, I think is extremely important, and maybe, at the health plan level it will be easier to get the data, but I think that that is a very high cost, very prevalent condition, and that we should continue to look at that, even though it's sort of stalled right now.

The infection rate, I recommend that

we make it more comprehensive to include the larger babies, to include GBS, because I think that infection is a source of considerable mortality and morbidity, not only in the low birth rate, but also in the larger babies.

And lastly, to Ashley's point, I was also disappointed in the structural attributes for maternal care, and I think, you know, I think that we should look at that and look at other states that are looking at that, looking at the ACOG, the ACOG recommendations, because I think that can tie in to the infant mortality rate and the maternal mortality rate that we're looking to find out more about.

And so, you know, I'm hoping we can fine tune that measure and come back with some that is more feasible and usable. Thank you.

MEMBER FLANAGAN: Just one quick comment. ICHOM, International Consortium for Health Outcomes Measurement. The contact is Stephanie Wissig. ICHOM.

International Consortium for Health

Outcomes Measurement. It's a very new bundle that literally is being finalized last month.

MEMBER AUSTIN: Their goal is to develop standardized measure sets for different health conditions that we, as an entire world, would be monitoring. And so they've taken on, I think, like 12 or 15 at this point different conditions.

DR. WINKLER: The other slide's on the portfolio. Okay. Go back. Back. There we go. Yes. Where we started.

So just to go through, any other thoughts on some of these other topic areas if you felt we were just focused? Naomi, what did you, did you have something you wanted to offer?

MEMBER SCHAPIRO: Yes. It's not, I'm not sure if there's anything like this is the database yet, but when we were talking about contraception and adolescence, there was not as much attention paid to the kind of, well it was the 15 to 18, I would even say 14 to 18 is, would be a really important age to look at.

And so if anybody's developing, and I think there's a lot of attention paid to like legal issues around confidentiality and what they are in different states, and there's a lot of attention paid to pregnancy rates for adolescence, and for pregnancy, delivery, abortion rates, but I'm not aware of a measure of an access to contraception, which is the preventive measure.

So I think, I'm not, and I'm not sure if anybody's developing it, although I'm going to sort of look into it.

But I would just encourage as to say that that's an important area, and that often in the way the data's collected for 15 to 21, we don't really get to see the folks who can't consent all the time.

DR. WINKLER: Okay. Well, obviously we've lost a lot of our colleagues, so I don't see any reason why we would need to keep going with this, but so thank you all very, very much for all the time you've put in, the

thoughtfulness, the discussions were great.

I apologize we've run over just a little bit. But the intensity of these conversations sometimes we end up doing that. Last words from our co-chairs, Nadine, Suzanne?

CO-CHAIR GREGORY: I would just say that this has really been an exciting process and you guys have contributed greatly and I really appreciate all of your input and I've learned from all of you.

DR. WINKLER: All right, so I think we're --

MS. THEBERGE: Thank you.

DR. WINKLER: We're adjourned.

(Whereupon, the above-entitled matter went off the record at 2:29 p.m.)

F

O

S

C E R T I F I C A T E

This is to certify that the foregoing transcript

In the matter of: PERINATAL AND REPRODUCTIVE HEALTH STANDING COMMITTEE

Before: NQF

Date: 05-03-16

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