NATIONAL QUALITY FORUM

+ + + + +

PERINATAL AND REPRODUCTIVE HEALTHCARE
ENDORSEMENT MAINTENANCE STEERING COMMITTEE

+ + + + +

WEDNESDAY
NOVEMBER 30, 2011

+ + + + +

The Steering Committee met at the National Quality Forum, Suite 900, 1030 15th Street, NW, Washington, DC, at 8:00 a.m., Laura Riley and Carol Sakala, Co-Chairs, presiding.

PRESENT:
LAURA RILEY, MD, Co-Chair
CAROL SAKALA, PhD, MSPH, Co-Chair
JOANNE ARMSTRONG, MD, MPH, Aetna
JENNIFER BAILIT, MetroHealth Medical Center
SCOTT BERNS, MD, MPH, FAAP, March of Dimes
JENNIFER BRANDENBURG, RN, MSN, Decatur Memorial Hospital
WILLIAM CALLAGHAN, MD, MPH, Centers for Disease Control and Prevention
KATE CHENOK, MBA, Pacific Business Group on Health
CHARLES DENK, PhD, New Jersey Department of Health and Senior Services
ELIZABETH DRYE, MD, SM, Yale School of Medicine
REBECCA GEE, MD, MPH, MS, Louisiana State University School of Public Health
ANDREA GELZER, MD, MS, FACP, AmeriHealth Mercy Family of Companies
CRAIG GILLIAM, BSMT, MT (ASCP), CIC, Arkansas Children's Hospital
KIMBERLY GREGORY, MD, MPH, Cedars-Sinai Medical Center

MAMBARAMBATH JALEEL, MD, University of Texas Southwestern Medical Center
BARBARA KELLY, MD, A.F. Williams Family Medicine Center
TERI KIEHN, MS, RNC, Intermountain Healthcare
MARYI SALGADY LESLIE, CNM, MSN, EdD(c), The George Washington University
NANCY LOWE, CNM, PhD, FACNM, FAAN, University of Colorado Denver
LEE PARTRIDGE, National Partnership for Women & Families
JOCHEN PROFIT, MD, MPH, Baylor College of Medicine
KATHLEEN RICE SIMPSON, PhD, RNC, FAAN, St. John's Mercy Health Care
SHARON SUTHERLAND, MD, Cleveland Clinic
ROBERT WATSON, MD, MMM, CPE, Baylor Andrews Women's Hospital
JANET YOUNG, MD, Carilion Health Systems

NQF STAFF:
HELEN BURSTIN, MD, MPH
JANET CORRIGAN
SHEILA CRAWFORD
EUGENE CUNNINGHAM
LAURA MILLER
SUZANNE THEBERGE

REVA WINKLER, MD, MPH
DONALD WASHINGTON

ALSO PRESENT:

JOSEPH CARPENTER, MS, Vermont Oxford Network (via telephone)

SEAN CURRIGAN, MPH, American Congress of Obstetricians and Gynecologists

MATT HOFFMAN, MD, Cristiana Care

JEFFREY HORBAR, MD, Vermont Oxford Network (via telephone)

ELLIOT MAIN, MD, California Department of Public Health

SUSAN MANN, MD, Beth Israel Deaconess Medical Center (via telephone)

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

JANET MURI, The Joint Commission

STEPHEN PRATT, MD, Beth Israel Deaconess Medical Center

PATRICK ROMANO, MD, University of California - Davis

MICHAEL ROSS, MD, MPH, The Joint Commission

ANN WATT, MBA, RHIA, The Joint Commission

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

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

8:05 a.m.

DR. WINKLER: Okay, good morning everyone. We need to go ahead and get started. Our agenda this morning is fairly packed.

And we do, we're aware that lots of you have 5:00, 5:30, whatever, flights and will probably be leaving, you know, threeish. We've planned for that, that's a very normal thing. So if you feel that's your circumstance, you're not alone.

So we do want to get started. We've put up a slide of the summary of what you did yesterday of the 15 measures reviewed. You rated nine of them as meeting the criteria for endorsements.

With that, we have another day of measure evaluation. And so, Laura, Carol, whoever.

CO-CHAIR RILEY: All righty, so first on the agenda will be Jennifer, to do

Incidence of Episiotomy.

And do we know if there is someone from Cristiana Care?

DR. WINKLER: Matt, you're here from Cristiana, right? And did we get the handheld mic for you?

CO-CHAIR RILEY: If not, you might want to sit right here on the side, next to a mic.

DR. BAILIT: So, this is the Incidence of Episiotomy. Just to orient everybody clinically, episiotomy is -- let me start that again. Sorry about that.

So just to orient everybody clinically to what a episiotomy is, to make sure we're all on the same page.

Episiotomy is when the doctor cuts the perineum right before delivery. Thought to facilitate delivery in some schools. Can also be used in an emergency situation, to try to hasten delivery.

The evidence against using it is

fairly significant in that it increases the chance of third and fourth degree tears.

Tears through the rectal sphincter or rectal mucosa, which are, needless to say, painful and can lead to incontinence of flatus and stool later in a woman's life.

So, that's the scientific background to this measure. This measure is calculated off of administrative data. And it is measured only in women who have a vaginal delivery.

It is the procedure incidence in women having a vaginal delivery over all vaginal deliveries. The one exclusion is shoulder dystocia. Generally considered an emergency, and would be an appropriate time to use an emergency procedure like an episiotomy.

In terms of what the work group thought at our conference call, the importance of measure was five high and four moderate.

Opportunities for improvement are quite high. It was six high and three

moderate. The evidence shows that there's a wide variety of rates in hospitals.

I don't have those numbers right in front of me. But it was a ten or 20 point spread between hospitals. And we also agreed, with one exception, that it meets importance.

With that, why don't we open up to the discussion of importance to measure?

CO-CHAIR RILEY: Any comments or questions? If not, we should vote.

Screen's over here. Everybody is trying to get oriented. The screen's over here, it's going to be a yes/no. Just as yesterday. Can't see when it opens up though.

CO-CHAIR RILEY: No. Can you scoot over to Matt's other side, maybe? There you go.

(Whereupon, the vote was taken, with 19 yes and zero no.)

DR. BAILIT: Okay, so that was pretty clear-cut. Move on to scientific -- so this is reliability and validity. In terms

of measuring, this is a procedure coded in CPT codes. And generally charged pretty well.

Procedures in general are coded well. Unlike other kinds of surgeries, though, Global Care charging in pregnancy, means they can't be charged for specifically. But still, I think the general consensus is that that coding is pretty reliable.

And there's been some work on whether the medical record as the gold standard, meets with the coding. That work is quoted in this document.

And it shows there's a pretty high fidelity. And that when there isn't fidelity it's equally divided between under-coding and over-coding. This is a random variation as opposed to systematic error.

DR. ARMSTRONG: With the coding issues, is it under- and over-coding the degree of the episiotomy, or is it with the laceration versus intentional?

DR. BAILIT: My understanding is,

because this is not an outcome, we're not looking at the, it's not the diagnoses of the degree. But it's actually the procedure. It's yes or no.

Any thoughts or questions?

DR. KELLY: Jennifer, is that captured as well in electronic coding?

DR. BAILIT: So this study, my understanding is that the electronic coding or the administrative billing data, which is what this is built off of. And that's the study of the electronic data.

It is -- what am I trying to say? There's fidelity there, with the exception of the few hospitals that don't, the few codes that aren't, don't have high fidelity, are equally distributed among over- and under-coded.

And if you want, I can try to find that number. I don't know if anybody else can find the line and verse faster than I can. Matt, do you remember off the top of your

head?

DR. HOFFMAN: So our validation study was done against NPIC. We had chosen a cohort of hospitals and looked at two separate Epics. Once again, we do use UB-92 billing information that compared it against the medical record. We asked that a subset of patients, both with and without episiotomy, be compared.

DR. WINKLER: Matt, could you turn your mic on?

DR. HOFFMAN: It is. I'm sorry, I'm just not speaking close enough.

DR. WINKLER: It is? Okay.

DR. HOFFMAN: Just to repeat myself. So the validation study that we did, involved NPIC, we used a cohort of the larger CWISH hospital, which is an entity underneath NPIC.

It is predicated upon billing data per se. But then we compared it to the medical record. We looked at two separate

Epics and basically found very good consistency between the medical record and the coding data.

In the majority of hospitals, where there was miscoding, it was a relatively small percentage. And once again, and Jen had mentioned, was equally divided, pretty much, between those who under-coded and over-coded.

DR. BAILIT: Other thoughts or concerns that you want to raise?

DR. DRYE: Question. Does coding affect billing in this case? Because we, in general, codes and claims data are more reliable if they're related to billing and audited.

DR. HOFFMAN: There is some individuation with the insurance plans, but for the majority of times it's captioned underneath global codes.

The other part if it's recorded in the medical record, is generally coders, hospital-based coders, who are choosing to do

this. And then generally they're fairly professional in their behavior.

DR. BAILIT: So global meaning that there's one charge, no matter what you do during the delivery?

DR. HOFFMAN: Right.

DR. SUTHERLAND: I can speak to this a little bit in my institution, because at the Cleveland Clinic, we started tracking this a few years back. And we found that there was a lot of documentation issue.

Many times this got documented in the nurse record, rather than in the provider record. So now that providers know it is being tracked, are taking responsibility for documenting more correctly. So that may help.

DR. WINKLER: I noted, Matt, on this submission for maintenance, this measure was previously endorsed three years ago. The level of analysis is just facility.

And I know three years ago the discussion was including clinician level as

the level of analysis. I'm wondering, was that deliberately excluded or -- because that is a change.

DR. HOFFMAN: So, from our perspective, we think the starting of the level of the facility is more appropriate in what caused general change rather than necessarily targeting providers.

I will say that I also participate on the AMA PCPI, and they are looking at it as an individual measure. So that may be forthcoming.

DR. WINKLER: Though, frankly, NQF's preference is to use a single measure with broadest applicability, rather than multiple little measures, as you'll see in our discussion of relating and competing measures later today.

DR. BAILIT: And just as a researcher comment, I mean, I think the confidence intervals around the individual providers are so not reliable, that I'm not

sure you could even look at it.

For any given provider, they're doing maybe handfuls. They're doing more than handfuls. That's concerning, but they have to have a lot of deliveries. So really the hospital level is a more stable number to look at, I think.

CO-CHAIR RILEY: Any further questions, or can we vote on scientific acceptability of the measure?

Nineteen yes, zero no.

DR. BAILIT: So is this understandable by the public? I think yes, and a lot of women going into labor, understand what a cut is. Most of them don't want it. Just the ick factor. It's publically understandable.

And then, is this meaningful, understandable and useful for quality improvement? I think so. I think there's fairly good consensus in the literature, that this is not a desirable thing. It should be

used sparingly.

And there's wide variation and performance gap, which would lead to the potential for quality improvement.

DR. PROFIT: My question just would be when the last three years, you know, do you have evidence or data that people have used this affect improvement, since it has been out there for a while now? So can we see some benefit of its use?

DR. HOFFMAN: So I can share, at least within my own institution, we've gone from a 16 percent rate to a four percent rate. So by simply tracking and reporting, when we do it on a physician level, because we have the ability off of our EMR to do that. We've been able to decrease our rates substantially.

I think California has had a like experience. I can let Elliot comment to that. But I know that it is being tracked in other institutions and have shown improvements.

DR. BAILIT: Matt, I'm just

curious. But does this tends to be a sort of, I may get myself in trouble here, but a procedure used by older providers, that used to be much more popular. As those folks are retiring, do we have any evidence that it's actually the measurement as apposed to just the natural attrition of that generation of OBs?

DR. WATSON: Since she just threw me under the bus.

(Laughter.)

DR. WATSON: I mean, I was trained to always do an episiotomy. You would much rather do an episiotomy than then let the pelvis shatter.

I'm still not convinced not doing an episiotomy is the right thing. But I don't do episiotomies any more.

But it's a big culture change. It's a big practice change, for many of the physicians, not just the older ones, but many of the younger ones as well.

And we have been tracking this on an individual basis in our hospital. We've had a dramatic effect. So, from a quality standpoint, it is, it has been very powerful. We dropped our rate, well, ACOG's rate, I think 2006, was 33 percent. It's gone down and down and down.

As a system with 22,000 deliveries, we were at about 18 or 19 percent. Now we're under ten percent, and going down.

So, there's a lot of change here. But I think it's a very powerful measure for quality improvement.

MS. BRANDENBURG: I'm curious, with your decrease in episiotomies, have you seen an increase in your tears?

DR. HOFFMAN: So, what one does see is a slight increase in anterior compartment tears. So, labial tears as a tradeoff. Those are usually minor, usually don't require stitching, fairly benign issues, compared to the other of third and fourth

degrees.

DR. BAILIT: Anybody else?

CO-CHAIR RILEY: Other questions, comments? We vote on --

DR. PROFIT: I guess this is question for the OBs. So, where's the tradeoff line between another? So one catastrophic tear to, you know, like the episiotomies. I don't know. What does it mean for an individual mom to have a grade four tear and incontinence? As opposed to many moms having an episiotomy?

How do we make that cut-off? You know, we can make the episiotomy rates go down. But, you know, how do we make that? Like maybe you can help us think about this.

DR. HOFFMAN: So let me clarify. So the issue is, by cutting an episiotomy, you in fact, weaken that tissue, predisposing a woman to a third and fourth degree.

So if you will, an episiotomy is on the pathway to raising the incidence of

third and fourth degree tears. So in fact it's a preventive measure against third and fourth degree lacerations.

DR. BAILIT: So, yes, lowering the episiotomy rate should lower the fourth and third degree tear rate.

DR. HOFFMAN: And that's been actually shown. There is an article in the French hospital, where they instituted a policy, and they demonstrated as they lowered their episiotomy rate, their third and fourth rate incidence went down.

DR. BAILIT: Just anecdotally, at our hospital, we almost never cut them. And we are actually in trouble with our residents, because they don't know how to repair fourth degree any more. They just don't see them anymore.

I don't want to throw my residents under the bus either, but, it becomes, you know, the senior resident case now, as opposed to what use to be a lower level case.

MS. BRANDENBURG: I guess I understand that. But I've also seen third and fourth degrees occur because an episiotomy was not done.

DR. HOFFMAN: So I think the argument has been made, initially, is an episiotomy easier to repair than a third and fourth degree that spontaneously develops? Yet if, you look at it, episiotomy clearly leads to more third and fourths across the incidence.

So, what you see is it's going to occur, predicated upon factors that remain beyond the obstetrician's control. The size of the baby, the size of the mother. There are actually ethnic differences as well.

So, no, we're not going to eliminate third and fourth degrees, but we will minimize them. And the one factor that is underneath the obstetrician's control, is the decision to cut an episiotomy.

DR. PROFIT: And wouldn't there be

a way for NPIC to track this more systematically, like across multiple institutions?

DR. HOFFMAN: So, in fact in our submission, we did do that. And demonstrated that, in fact, what we see is this wide variation within hospitals. So one of our hospitals had a 33 percent rate and another one had, I believe it was a less than ten percent rate.

So there is tremendous variation. The answer to the earlier question, we are seeing a decrease with time, as we're seeing a cultural change. Yet it still remains a significant percentage of women.

And particular hospitals that are having high rates of episiotomy. As attention has been focused on it, we're seeing decreases in the overall rates. We're thinking that's leading to improved outcomes.

CO-CHAIR RILEY: So can we go ahead and vote on usability for this measure?

Remember here we are doing high, moderate, low.

DR. DRYE: Can I just clarify? So, someone said you didn't expect it to be used in public reporting. But really, we're voting on whether it is suitable for public reporting across hospitals.

DR. BAILIT: I never -- at least if you are referring to me. It wasn't that hospital level, it was physician level I'm not sure is ready for prime time. To the hospital level, absolutely. In my mind at least.

CO-CHAIR RILEY: Can everybody press it one more time, so that we don't get the change?

(Whereupon, the vote was taken, with 14 high, 4 moderate, and one low.)

CO-CHAIR RILEY: All right, so moving on.

DR. LOWE: Yes. This is just a request. Could I ask that we use more inclusive language, acknowledging the fact

that not all births are attended by physicians? Many of them are attended by nurse midwives.

DR. BAILIT: The next slide up? Can we have the next? Yes. So in terms of feasibility, these are data generated during care. It's routine administrative data. It is electronically coded.

And we talked a little bit before about the fidelity of the coding, compared to the medical record. So I think, as Matt and his group has shown, that this data is fairly easy to use and manipulate across a wide group of hospitals and able to make comparisons.

CO-CHAIR RILEY: Can we vote on feasibility?

DR. WINKLER: Fifteen high, five moderate, zero low. I don't know if anybody's on the line or not.

CO-CHAIR RILEY: And moving on. So overall suitability for endorsement, for this measure. So 19 yes, 1 no. Thank you.

So our next measure is Exclusive Breastfeeding at Hospital Discharge.

DR. ARMSTRONG: I have that one.

DR. WINKLER: Can I just say a few words? The next measures we're going to be introducing are from the Joint Commission. These four measures were endorsed during NQF's previous work. And they had different measure developers.

Since that time the Joint Commission has partnered with those original measure developers to bring these measures forward under the stewardship of the Joint Commission. All right?

So they're expected to, you know, they've made this a little bit of evolution and transition.

DR. ARMSTRONG: Okay, right. So this is exclusive breastfeeding during hospitalization. It's for singleton term infants. I think, unlike episiotomy, it's pretty self-evident what exclusive

breastfeeding is.

You know, multiple recommendations from multiple medical professional societies and the Department of Health and Human Services and the Healthy People 2010 and 2020 goals.

Lots of medical benefits for the baby and the mom that have been described. And significant gaps in performance, relative to the Healthy People 2010 goals.

For exclusive breastfeeding at discharge, the goal is 75 percent. Data from the Joint Commission, looking at four quarters of data reported in 2011, have a rate at 41.5 percent. So, little more than half of where we're supposed to be.

The idea of exclusive breastfeeding at discharge is that it increases the likelihood of continued breastfeeding at three months and six months.

Some data presented that for women who exclusively breastfeed at discharge from

the hospital, they are seven times more likely to be breastfeeding at one month, compared to women who have some level of breastfeeding, but it's not exclusive at discharge from the hospital.

There's also data presented on disparities by race and ethnicity. So from a importance point of view, you know, our group thought, obviously, it was very high. Both the quantity, quality and consistency of the data is high.

CO-CHAIR RILEY: Questions or comments?

DR. BAILIT: I guess my concern with this one is that it's very susceptible to who your patients are, as opposed to what the doctor is doing.

And I know that we talk about racial disparities. We don't want to excuse those, we don't want to excuse the economic disparities that exist.

But, that, to a large degree those

are outside of the physician's purview and ability to change.

If you've got a 15-year-old mom with no breastfeeding support at home, you can maybe convince her to do some things in the hospital. But that chance of actually having it go forward is little.

My other concern is that the hospitals who are trying so hard do this go in the other direction of patient autonomy.

So that they're forcing people to breastfeed who -- there are, you know, moms who are quite sick, who really just need to rest, are being forced to breastfeed as well.

So, there's a tradeoff here, and I just want to make sure that we're not incentivizing people to go in the other direction.

DR. LOWE: Yes, I think we're little bit talking about apples and oranges. Because, first of all, it's much more than the provider who has an impact on whether the

mother goes home exclusively breastfeeding.

A lot of this has to do with nursing care. And if we don't acknowledge the fact that it's the whole system, that's what healthy, or, excuse me, Baby-Friendly Hospital Initiative is all about. Creating that environment that truly supports breastfeeding from the get-go.

One of my concerns about this measure is that it's unrelated to intention. The mother who comes in intending to breastfeed is not in the equation anywhere.

Perhaps that might be another measure that we could talk about. But, I really believe, and I think there's ample data to support, that if the environment supports the mother who intends to breastfeed, she will be successful.

And I really highly endorse this measure. And one is to think beyond, it's not about the provider. It's the whole environment of the hospital or the site of the

birthing, of whether or not the mother goes home exclusively breastfeeding.

DR. BAILIT: I would agree would agree with that. I guess my concerns still is" that's true for the mother who intends to breastfeed. For the mother who's dead set against it --

DR. LOWE: That's not what this measure is about. It's about who goes home exclusively breastfeeding.

DR. BAILIT: Right, but for the mother who doesn't want to breastfeed, who -- the hospital is going to measure on the quality. And that mother's going to be a ding. She gets tremendous pressure to breastfeed.

DR. LOWE: I would say there's other evidence from the nursing literature, that if even mothers who we consider a high risk, who we consider unlikely to breastfeed, if they're appropriately taught and brought to understand the benefits of breastfeeding.

Any mother, regardless of situation, wants to do the best thing for her baby. And I've helped plenty of young mothers breastfeed very successfully. I've helped plenty of disadvantaged mothers.

And I think there's evidence in the nursing literature, particularly from Paula Meier at Rush University, that shows that supporting, even in the NICU supporting very disadvantaged mothers to breastfeed can be highly successful. And often what they will say is: nobody told me how important it was for my baby that I breastfeed.

DR. BAILIT: I guess in my last set of rebuttal, it would be that we're seeing a tripling of our methadone moms. And our moms in polypharmacy, and our pediatricians don't want those moms breastfeeding.

For I think good reasons, and there's no exclusion for that kind of thing. So, you know, depending in the hospital, that is very different than a private community

suburban hospital, when you have a lot of methadone moms and a lot of poly-drug use.

So, I just think we need to build some of those medical exceptions or at least an acknowledgment of some of the extenuating circumstances into this.

DR. GEE: Louisiana has, we're 50th in the nation in breastfeeding. We have tremendous disparities. Our African-American women initiate at much, much lower rates.

We're talking 20 percent, 30 percent in some settings. So let me just start by saying, this measure, I think, is fine. But it sets the bar extremely high.

And so my comments later today, when we talk about new measures, we need new measures for breastfeeding. I think this is, if you're in a setting where breastfeeding rates are high, exclusive breastfeeding is fine to look at.

In Louisiana, our rates are practically zero, of this rate. You look at

circumcision, there's some concern at Women's Hospital, which is our largest birthing hospital, if you give sugar water, is that, then does that mean you didn't exclusively breastfeed?

So there's confusion about the measure. It's very difficult to measure, because you have to -- it's every single episode of care and feeding of that infant in the hospital being exclusive.

Rather than looking at, what we're thinking about putting in our report card is breastfeeding at discharge. Because at least you're then just looking at one day of data, instead of the entire hospital stay.

And discharge, if you're breastfeeding that day, that's a good indication that you're intending to breastfeed at home.

Another thing, I just looked at our PRAMS data, our population-based data of pregnant women. About disparities in hospital

experience, once women intend to breast feed, we found that African-American women had much less hospital support.

Even once they had initiated, so they were more likely to get a gift bag. They were less likely to get numbers in counseling and lactation consultant referrals in hospitals than white women, even if they had both wanted to breastfeed.

So I think there are issues associated with disparities here that needs to be looked at in metrics.

But my issue with this measure particularly, is that it's so labor-intensive to collect it. Hospitals complain a lot about it, because it's every single episode of feeding.

And we're not looking at disparities. And we are ought to maybe look, if we want to do improvement, we ought to look at something that we can improve easier, because we are so lagging.

States that really need help, we ought to be looking at a lower mark, maybe breastfeeding at discharge. So this, maybe we could we talk about that more later.

DR. ARMSTRONG: Yes, the question there actually, it didn't come up in our discussion group. But for hospitals that are reporting on this, are they looking at every episode of breastfeeding, whether that is exclusive? Or is it intention, an early note in the chart of intention to exclusively breastfeed?

DR. GEE: In Louisiana, it's measure every single episode. And so it's a lot of work.

DR. WATSON: Just to echo what Jennifer was saying, in our area we've had some mothers complain to the hospital, or even change hospitals to deliver their babies, because they had no intention of breastfeeding.

Perhaps they tried it last time

and it was unsuccessful. And they were so belittled and intimidated by our nursing staff, that was so focused on this measure to be Baby-Friendly, that they did not want to come back.

And so we're actually having some of our patients that felt that way come to our Baby-Friendly committee next month and talk about their experiences.

So I think it, I think there can be too much emphasis on the ones that don't want to breastfeed, for one reason or another.

DR. WATSON: Hi, I'm sorry, did you want to go?

CO-CHAIR RILEY: Kathleen, please.

DR. SIMPSON: I just wanted to say this measure has been very helpful at our hospital. We are a large community hospital, so not like what you have, we're more like Women's Hospital, I guess.

Our pediatricians were routinely ordering sugar water after every birth,

supplementation in the hospital. And when this came out, this was extremely helpful. It changed practice right away.

And yes, it is labor-intensive. But it has been very positive. We were able to get more lactation consultants on staff because of it.

We were able to change those quirky practices that had not been able to be changed for evidence, but they were changed by now we were looking at this. And it's been very helpful.

We do measure every feeding to make sure that the baby didn't get any sugar water or supplementation. So it does take a lot of time, but it's been very positive.

DR. DENK: Hi. In New Jersey we've, you know, we actually do risk adjust for breastfeeding. We've been reporting breastfeeding at discharge for, you know, 15 years in our hospitals.

And they weren't really paying

much attention, because they would always argue that our, you know, populations differ between hospitals serving Newark and hospitals serving Princeton.

And there was just never going to be any kind of, you know, dialogue on that level.

So we started doing risk adjustment. We also, you know, started bringing out evidence which was, you know, I mean, CDC did a great job. The developers, by the way, did a great job summarizing the whole literature on what is the hospital's piece of all of this.

And that awareness has only come around in the last four or five years, that you have documentation that providers will pay attention to, that say, you know, there's a hospital piece.

PRAMS, we've used PRAMS to actually demonstrate that women who start breastfeeding, but then are not exclusively

breast-fed during the duration of hospital stay have different outcomes, you know, eight months, eight weeks post-partum.

So, you know, the face validity of this with providers is tough, if there isn't some kind of an acknowledgment that some moms are different.

And, you know, our risk adjustment is basically on the base of race and educational attainment. And that's all, so it's easy to do. But if you don't want to do it, you could stratify by education and race group. And maybe that, you know.

So that's something at least, at a minimum to recommend sort of stratifying. So to increase the usability of this with the population you most need to talk to. Which is the, well, one of the populations, which is the provider population.

Oh, by the way, we did a study that said that about 40 percent of the variance in hospital to hospital,

breastfeeding rates came out via the risk adjustment model.

So it's not at all insignificant, although it isn't the whole story. And that's the whole point, is that it's got two pieces.

CO-CHAIR RILEY: Yes.

DR. JALEEL: In terms of Jennifer's last comment about the provider not willing to do this. One of the denominator exclusions is documented reason for not exclusively breastfeeding. So there is a documented reason then.

DR. BAILIT: Depending on where you are, though, does HIV count, does methadone count, does methadone with cocaine count? I mean, it gets, how you parse that gets tricky.

DR. ARMSTRONG: So it sort of moves us to the next area of scientific acceptability. So the exclusions are, again, it's --

CO-CHAIR RILEY: Actually, are we

done with importance, like is everybody is on board here? Because maybe we should vote on this, and then move to the next thing.

Absolutely, feel free.

DR. MAIN: What we've learned is that there's no such thing as an easy measure here. No matter what you do, there all kinds of what. Something as simple as breastfeeding gets very complicated when you try to turn it into a measure.

I think there's a couple of key points here. First is: the disparity issue was a big issue in California, because it's a huge state with a lot of different ethnic groups. This is an issue we've been working on for five or six years in California. It has shown that actually if you put a whole community at work, you can actually change ethnic acceptance of breastfeeding quite successfully.

We have some counties that have taken this on through their MCH programs very

aggressively. A lot of public outreach that have markedly increased the exclusive breastfeeding rates in hospitals in southern California, where the rates are much lower.

Also, there are hospitals that serve minority groups, in different parts of the state that do much better than other parts of the state.

In Oakland, for example, which has a very high African-American population. They have very good breastfeeding rates, whereas other parts, other hospitals in other parts of the states don't do as well.

So there are opportunities to overcome the ethnic differences. But we're not looking for 100 percent. That's really the whole take-home thing.

And it gets to the point of over-zealous nurses. Again, there has to be some sensitivity training, that we're not looking for every mom to do exclusive breastfeeding for the whole stay.

We're looking for, you know, a reasonable number, and if the mother is not interested that should be very acceptable. But that's a slippery slope, if you start putting that into the measure definition, though.

And, again, if you accept less than 100 percent, that should come out. There are a lot of documented reasons that would exclude you from breastfeeding, including methadone, including HIV. Those were all on the Joint Commission list now.

Those end up not being a big number, though. In most hospitals, you know. Maybe Jen, at Metro it's a little higher number. But in most hospitals it's still five, ten percent, at the most.

There are ways of doing the documentation that make it less onerous to collect. Maybe you want to talk about that.

MS. MILTON: Actually, we're out doing reliability visits right now on this

measure. And depending on how you have your charts set up, if you go to the feeding record you can get all of this data.

And most newborns that are healthy aren't going to be in the hospital more than a couple of days. So in my experience, probably at seven or eight different hospitals, it's probably been less than five minutes per record, to take a look at that information.

If they're admitted to a NICU, they're immediately excluded, so you don't look further in the record. And then you strictly look at the feeding records.

If they've had strictly formula, then you would have to look further to see if there was an indicated reason that's acceptable in our exclusions.

But if they're a combination, that tells me that there wasn't a reason. You wouldn't start breastfeeding if there was a maternal reason to exclude.

So it really hasn't been too difficult to find this information in the records.

DR. MAIN: We looked for a number of years, trying to compare women who exclusively breast-fed versus women who breast and bottle fed. Combined feeding.

And it was sort of interesting, most hospitals claimed extraordinarily high rates of breast and bottle, 90-plus percent. But where you saw the biggest discrimination between rates, or the biggest variance between rates, was in the rates of exclusive breastfeeding. And that's what's been associated, in a number of pediatrics studies, with long term success of breastfeeding at the three and six month mark. Which is really -- what we're doing is a surrogate for three and six month breastfeeding. Because that's almost impossible to collect with current date of collection systems.

So, we didn't feel that combined

breastfeeding was a reasonable target that would show value. Because it could be just a baby at a breast briefly and mostly bottle feeding would qualify under that. And so, we really wanted to shift toward exclusive, knowing that it would not be a 100 percent, though we have hospitals in California right now, ranging from 40 to about 85 to 90 percent. And that's actually up about 20 percent from where it was a few years ago. Both on the low end and the high end.

DR. GELZER: Our health plans serve exclusively vulnerable populations. And I just think that this measure is hugely important to those populations.

And, sure, a safety net hospitals may not score as well on this measure as is other, you know, community hospitals and nice suburbs. Having said that, Medicaid health plans are measured prenatal, post-natal visits, just as commercial plans are, and we don't score as well, either. But that doesn't

mean we shouldn't measure the measure.

DR. DRYE: Have you thought about stratifying for race? I'm over here. Because I just think in this measure. I mean, we need to always be really cautious about how we think about disparities. And you don't want to set different goals or benchmarks.

But it might be informative to hospitals to see that some people are doing very well -- some hospital systems are doing well with some populations. And they may not be able to make that direct comparison without stratified results.

MS. MILTON: The issue of race is very difficult, as far as defining it as a reliable data element. And that's been the problem. If it was strictly you had two African Americans, well the baby would be African American. But what if the mom's half African American, half Hispanic, the father is Chinese? I'm just throwing that out there.

Seriously, and that's been the

problem. We'd like to be able to do that. But until we can find a reliable way of measuring race it's almost impossible to do that in a reliable manner. Where it would be meaningful.

MS. PARTRIDGE: I really think that particular argument ought to go away. I live in the District of Columbia and we use to have arguments about whether a baby was Hispanic or African American, you know. We've now got rules and I think most everybody building measures follows them.

We haven't talked at all about the benefit to the child, of the breastfeeding. And this happens to be a measure that the National Partnership and, I personally am very supportive of.

As a working mom, many, many years ago, my OB suggested that I try breastfeeding for a month. Because four weeks was all I had of leave. And, of course once you've done it for a month, it's so easy, you're just keep

going. So in fact it was very, very helpful. And we really hope that this will continue to be in the measure set.

DR. PROFIT: I think what you said there was really interesting and it strikes me that it's almost like a social policy issue that we're attacking at the wrong end.

So, it's like we're measuring the providers with something that employers aren't allowing, you know, mothers time for it. I'll just put that out there.

MS. PARTRIDGE: That's actually one of the goals that the Partnership hopes to work on. Work place friendly for the nursing mom.

DR. BAILIT: And I would just echo that. I think this is a great population health measure. I just don't think this is a great health system measure.

I'm very intrigued by the fact that you said you could raise it in the community. If we wanted to collect this at

hospitals and report it at the community level, that would be great. Because then you're really bringing into play that bigger support system, it's not just the hospital.

DR. MAIN: I think what we've found, though, is that there's a huge amount of impact that the nurses and the physicians have in hospitals on the mother's actual follow through in activity on this area.

So, it's -- it's really about systems change. Both systems in the communities as well as systems in the hospital and the obstetric and pediatric community.

Yes, it needs some longer term change in the employment. But this is really focused on the parts that we can control in the health care system.

DR. GREGORY: I would just echo that. That because of this measure, hospitals are now putting into place, lactation support. And it's trickling down, you don't need just a lactation specialist. All the nurses are

now being trained.

Where the big fall out is in the middle of the night. That's when you need the most support. And I have to say this is one time where it's not the providers, but it's the doctors that needed to be educated. Because, even though we know it's a good thing, you know, the doctors, the obstetricans were saying, oh no, she needs her rest.

And so, it really took, you know, looking at the evidence and bringing it forward, and have pediatricians nagging the obstetricans to really help make this system change.

And I actually think sometimes the education was a detriment that, you know, that the more educated you are, the less you thought it was a big deal. And so, bringing this evidence out and sharing it with everyone actually led to some pretty significant system changes.

DR. ARMSTRONG: I would also

comment that some of the healthcare reform legislation that's been enacted provides access to lactation consultants and breast pumps, which will help women, working women, particularly continue longer term.

So, some of the system changes are sort of moving in the direction of supporting this more for more women.

DR. PROFIT: Just with the race debate that we've had before. And I'd appreciate our colleagues from California's input on that.

We know that some of our Asian population, they end up breastfeeding at home, but for cultural reasons they do not want to do it in the hospital. Even when we talk to them, support them, said, you know, we really would support you doing it. They just don't want to do it in the hospital. But they will breast feed as soon as they are at home.

I don't know if you've had any experience with that, with the larger Asian

population in California.

DR. MAIN: My hospital delivers 6,000 babies a year, of which 40 percent are Asians. Probably the largest Asian delivery service in the country. That is a population that has less exclusive breastfeeding in general. But our hospitals as a whole still has 75 percent overall exclusive breastfeeding during hospital stay.

Again, it's a matter of communications and working with the community ahead of time. And you're looking for shifts overtime.

There are, certainly, ethnic variations to start with. California does report this, both your hospital total exclusive, and we do break it down by race. Using birth certificate race categories. That is a reason --

DR. ARMSTRONG: And the data in this report suggest that for Asian women who are exclusively breastfeeding at discharge,

about 60 -- 59 percent of those are breastfeeding at one month, versus only 20 percent who leave the hospital not exclusively breastfeeding.

CO-CHAIR SAKALA: Several quick comments to Lee's comment about benefits for the baby. I'd like to say this is a great measure for moms, as well. So I think that's very valuable.

I'd like to echo the sediment that the fact of the disparities mean that we could really benefit from really strong measures.

And finally I want to mention mPINC survey from CDC, which shows phenomenal variation in hospital practice, and around practices around breastfeeding and tremendous opportunities for improvement.

CO-CHAIR RILEY: Okay, so I think we should move on. Can we vote on importance to measure and report for this measure?

CO-CHAIR RILEY: Can you put your phone on mute, please? If you're on the line?

Thank you.

DR. WINKLER: Twenty-one yes, three no.

CO-CHAIR RILEY: So let's move on.

DR. ARMSTRONG: Okay. So scientific acceptability, we discussed some of these already, just the exclusions in the measure are multiple births, babies admitted to the NICU, and experience of death. A baby enrolled in a clinical trial, or mother enrolled in clinical trial length, stayed greater than the 120 days. Transferred to another hospital. Infants on TPN. And documented reasons to not exclusively breastfeed, which are specified in the measure. They include HIV, illicit drug use, mother taking other drugs, chemotherapy agents, et cetera. TB, active breast herpes. That's really it.

This question about whether there other sort of reasons to breastfeed that are being gamed in the chart. I think, Jennifer,

you sort of suggested maybe that's going on. There's not really a discussion about that.

DR. BERNS: So, I have two questions. One is general, because it comes up in a number of different submissions. And then the other is more specific, I think.

First, reliability testing, you mentioned that you're going to be doing that, you know, this year. And, I'm just wondering in general how long it takes to do reliability testing?

Because I noticed on a couple of these measures that were approved in 2008 that were about to do reliability testing. What's the sort of time span on that? I've noticed this in a couple of different measures that we're viewing today.

MS. MILTON: We're actually doing reliability site visits on all five of our measures in the measures set. And we started in October and we'll be finished in January, as far as the site visits.

Our statistician selected 12 sites randomly across the country. And we're re-abstracting approximately 30 medical records per site.

DR. BERNS: So, let me restate my question. So, for a measure that's approved in October 2008. Now, I assume because the package wasn't implemented until 2010 -- I mean, I'm making some assumptions, I'm just trying to figure out what's the -- it just seems like -- I wish we had the data now. Is what I'm asking.

DR. ARMSTRONG: You do. So the report says that, at least in four quarters that they looked at, that the rate of exclusive hospital breastfeeding changed from 39.7 percent to 42.9 percent. So some change in a year.

MS. MILTON: It was a matter -- it took time. Once they were endorsed for us to actually specify them, to get our manual out, to give everybody six months to embed the

measures. There are performance measurement systems vendors, who acts as our intermediary between the hospital and the Joint Commission. We had to review all of their data collection tools to make sure that they were, you know, valid and reliable.

So there's a period of time, kind of a set up time. And then actually after we received the first couple of quarters of data, we determined that there were additional enhancements that were needed to the measure specifications.

So we're actually doing reliability visits on our second version of the manual. It's not the first version. And actually we're on the third version now.

There's always a lag time. You just don't -- you'd like to go out sooner, rather than later. But we had to really wait until we felt we had something that would be meaningful to review, as far as medical records.

DR. BERNS: And just to follow up on the validity testing, you do have, and this is great, you know, the automatic, the automated feedback system. And you had, at least in this report, you had 130 submissions. And you had some interesting comments here, particularly regarding the denominator.

But -- so my question is, we heard some anecdotal reports here regarding the difficulty with the measure. I'm wondering if any of those 130 submissions, if you've heard similar responses, in terms of the challenges of this measure?

MS. MILTON: Difficulty as far collecting the information, or how it's constructed? I'm not sure what your question is.

DR. BERNS: Difficulty with measuring. I mean, we heard from Louisiana and -- I'm just curious.

CO-CHAIR RILEY: The fact that you're collecting every single --

MS. MILTON: Okay. Well, actually we really haven't in the automated system. I'm just giving you my experience having gone out to, probably, what I'd say, eight different sites now. Each one with a different type of a medical record. Some of them are EMRs, some of them are paper-based, some of then are hybrid, where they scan it in.

And, really, the bottom line is, especially for this measure, you go to the feeding records. That's where you look. And if the baby's there a day or two, you're looking at a couple of days worth of feeding records.

And if the flow sheets are well constructed, it's pretty simple to see. Because usually most of them will have a line for breastfeeding, a line for formula feeding. If they have a bottle, what was in the bottle. Because sometimes we do allow -- they don't have to actually be fed on the breast. It's

the breast milk that they received that we're looking for.

So if they've clearly documented in the flow sheet, I'm not spending more than five minutes on a record, for the most part. It's a pretty straight forward review of the record. It's not like another project where I looked at blood management, where if they might have had a dozen transfusions, that's a lot harder, more labor intensive. This is fairly straight forward.

DR. BERNS: I think one of the key systems changes to make it easy for hospitals to collect this measure, is to actually to tweak your forms, to make it real clear how you're suppose to chart. And where to look to find the data. Once you do that, things fall into place.

CO-CHAIR RILEY: So --

MS. MILTON: We also do sample for this measure, so we are not looking at every single medical record, as well. So it's based

on the number of discharges that you have, or deliveries per month. Then you would sample accordingly.

CO-CHAIR RILEY: Okay, can we vote then on scientific acceptability of this measure? So reliability and validity.

DR. WINKLER: Twenty-two yes, two no.

CO-CHAIR RILEY: Okay, so we can move on.

DR. ARMSTRONG: Okay. Next, for usability, our group thought that the measure was very easy to understand. Good for public reporting, good for quality improvement efforts. Public could easily understand what it is.

And there's some data presented that it has actually made a small difference in one year of reporting on the measure. Comments?

CO-CHAIR RILEY: Okay, can we vote on usability?

DR. WINKLER: Sixteen high, six moderate and two low.

DR. ARMSTRONG: Okay, and then from a feasibility point of view, we've already discussed the ease of collecting the data. I guess there's some, perhaps, difference of opinion about it. But our group thought generally that it was -- the measure was rated high on feasibility. Four highs and one moderate.

CO-CHAIR RILEY: For those of you who weren't looking at the screen, I wasn't either.

DR. WINKLER: High nine, moderate 12, low three.

CO-CHAIR RILEY: And then overall suitability for endorsement.

DR. WINKLER: Twenty yes, four no.

DR. PROFIT: I would just mention that maybe a similar measure could be constructed for the NICU and is already in use, you know, in many consortia. It

shouldn't be exclusive milk feeding, of course, but maybe any human milk at discharge. It was a very highly rated measure by many of the expert panel that we conducted.

DR. JALEEL: Just want to comment on the document itself. Having gone through many of these documents, I think this has been very well put together. And I just want to comment the group. And this can be put as an example of what the documents should be like.

DR. DENK: Hi, I just wanted to make one comment if it's alright. As a person who makes, who does report cards and other kind of data reporting at a state level -- and I don't have access to charts. I've got to rely on instruments, like the electronic birth registration system. I just like the Joint Commission to think a little bit about how to harmonize with that.

I mean, maybe there should be two versions of the measure. Because, I'm not sure, but I think the 2003 standard is

breastfeeding and -- exclusive breastfeeding at discharge. Not exclusive breastfeeding through out the hospitalization. I can't -- if anybody knows better, I'd appreciate that. But, you know, not everybody who's doing these measures and loves them, is doing them, you know, has direct access to charts.

DR. GEE: I wanted to add to what Chuck said. You're, Elliot, probably talking about hospitals that are motivated to support breastfeeding. When you're trying to change behavior in hospitals that aren't motivated to support breastfeeding. And you've got -- as difficult as this is to collect, if you're not setting your system up that way -- it's not -- and you can choose the measure or not.

We need to do measures that are also more processed based. Looking at -- you know, we have Baby-Friendly Hospitals, that requires a certain architecture of your labor floor.

In Louisiana we don't have a

single baby-Friendly hospital, for example. And for some states that are trying to improve, and you're looking at grading them, we've had a heck of a time of what to put on our report card.

We have something called the GIFT Certification Program that varies by each state, which is getting rid of the formula bag. But I think we can talk about this later, too, but Elliot, you may be leaving.

What other types of intermediate process measures have you looked at? Because from the state perspective, and the worst performers' perspective, where you want to change behavior quickly, we really just don't have anything that's validated to grade them on. And so we're creating our own as we go.

DR. MAIN: We started out years ago using the data that was reported on the newborn screening forms, which is a quick but very dirty form of data, because it often records what the intention was, rather than

what the reality was.

But that's a start, as a bridge to get to a more refined measure. But, again, I think the intent here is that this is something that focuses attention on this issue and makes the hospitals confront it, you know, in a way that can be creative.

DR. GEE: If the measure were required it would be a whole different ball game. And then we could motivate behavior. But for the non-motivated we are using newborn screening for our report card. But it's just a single time measure. And then discharge, it sounds like New Jersey's using.

CO-CHAIR RILEY: Okay, so we're going to move on. And our next measure is Elective Delivery Prior to 39 Completed Weeks Gestation.

DR. GREGORY: All right, is this good to see that?

CO-CHAIR RILEY: Well, not behind Elliot, you can't see it.

DR. GREGORY: Okay, so this is elective delivery prior to 39 completed weeks, which is 37 to 39 weeks. And this was a very interesting discussion. Pretty much the group felt this had high impact because of high burden -- all right, yes, I will. Anyway, I'll do it from memory.

We felt that this was high impact, because of the high burden and the relatively -- based on several publications, the relatively high number of elective deliveries that were occurring at this time period.

We felt that there was a high opportunity for improvement. And so it met the importance goal. There was pretty much consensus around that. And the evidence to support it was pretty good.

I think the biggest bugaboo in this indicator was, one, there were very generous exclusion criteria. But even in the face of -- some of which may not necessarily have been appropriate, but they were very

generous in their exclusion criteria. But, then there were many criteria that should be excluded, that aren't excluded, because there are no codes for them.

And the one that comes in mind, right away, is classical C-section. That is a very good reason to have a delivery earlier than 39 weeks. But there's no specific ICD-9 code for it.

But in general there was pretty good consensus about that. So I'll open it up for comments.

DR. DENK: Yes, I wanted to comment just a little bit on the exclusions. You know, New Jersey's been working on this for about a year and a half. We've saw the 39 week thing as a low hanging fruit. And we got half the hospital chairs in the state, you know, OB chairs together to try to get them to sort of jointly agree to a ban. It wound up that hospitals had to do it, you know, voluntarily and one on one.

And some hospitals have really hard stops, where if you don't have medical documentation you can't get scheduled, you can't get admitted, you can't get nursing staff to run the induction and so on.

But when a hospital puts a hard stop in like that, then there's response from physicians, OB physicians. One -- I was doing CMEs and one deputy chair who was responsible for screening all of these cases said that in the week after they put a hard stop policy in he had five recorded cases of oligo in the 38th week that had to be induced, right?

And oligo is here on the list and I was just told, I'm not the provider, this isn't my training, but I'm told that oligo is a pretty unreliable measure. And if you do it a couple of times over the course of two days you can get any result you want, right?

CO-CHAIR RILEY: Or a dead baby if you wait long enough.

DR. DENK: Okay. That's been the

New Jersey experience that, you know, that you have to be really careful about these.

DR. BAILIT: We looked at the Ohio Perinatal Quality Collaborative that brought our rate down from around -- we decreased it by 60 percent over a year. And we looked at diagnosis creep and this data has been published in abstract form but has not yet been published in paper form.

The elective indications went down, but the indicateds did not go up. Suggesting that diagnosis creep is not a significant player.

At some point you've got to trust people to document the right thing. I'm not saying that games don't get played, but I'm just saying on a big picture it's not happening. At least in Ohio where we looked at it.

DR. GEE: Chuck -- we had a similar process to what Chuck did in New Jersey, where last year we challenged our

hospitals to be 39-week hospitals. We now have 85 percent of our hospitals are 39-week or above hospitals.

But out major problem has been measurement. And this measure is difficult and states -- in many states, Medicaid does not require ICD-9 modifiers and so we only have, let's say two per admission per billing.

And so we could not -- this would require chart abstraction if we were for every case or every single chart, for complicated medical reasons the chart abstractors may not understand.

And so the way we've gone is we're revising our vital records system and creating a check list for each physician delivering that notes the indicative reason. And we've created a consensus panel around what the reasons are. And they were actually more restrictive than these. Again, the classical c-section is an important one that every one agrees, that we've talked to, and I agree with

as a reason why you would deliver early.

But oligo and gestational, depending on suspected macrosomia, some of these are softer indications. But we've added them, they're still on our checklist.

But it was important for us to go through this process, I think ACOG needs to be a bigger leader in terms of making stronger statements on 39 weeks and helping providers navigate these decisions.

But what about the issue of not having the ICD-9 modifications that would really clearly indicate what these are? Elliot's actually worked with us trying to help us. But we realized it would be hopeless to try to do it from a billing standpoint and we had to go to our vital records system.

So it is a challenging measure in that regard. And then people will always question the classical -- the absence of that modifier as a very important exception that's

missing.

DR. ARMSTRONG: I would comment that we try to build in financial incentives for both physicians and hospitals around this measure but, again, you can't see it in the code system. So you can't independently verify it. So for 2000 ICD-10 is there greater specificity in the coding?

MS. MILTON: Yes and no. There's a lot more codes in ICD-10, but like the prior classical cesarean section, it's not going to be in the first iteration of ICD-10 codes but I do understand it will be in the next release.

So eventually that will be something that can be coded, as will prior myomectomy, that's another one that we hear about as well. But that will be the iteration after the second iteration, where we'll see that.

So -- I'm trying to think of some of the other things that we've heard. It's a

lot more specific, like if you're doing a procedure, like the right side versus the left side, versus lateral versus medial. There's a lot of that when you're looking at procedures but a lot the diagnosis codes, not really.

DR. ARMSTRONG: So does this group ever advocate to the AMA that there be specific coding around elective deliveries and gestational age of interest?

MS. MILTON: Gestational age is another one, I believe, right, Sean? And parity, correct? Okay, those two are coming eventually. And then parity is in the next iteration. So it's coming, but it's not here yet.

DR. ARMSTRONG: Do you know what year we're talking about?

(Off microphone comments.)

DR. GEE: And some states, the state of Texas this year took the tack that they would do Medicaid modifiers for billing

that indicated whether the delivery was medically indicated or not.

My problem with that, obviously, is there would be a lot of bias in the coding, if you're voluntarily stating that you're not going to be paid on your form.

But my understanding is that no state has been able to get around the issue of ICD-9 coders. Because you're not going to get paid based on whether she -- you know, once you go down a very long list of exclusions, that you're not getting paid based on it, there's very little motivation for coders to put those codes in.

MS. KIEHN: I'm from Intermountain Healthcare, for those of you that don't understand them, we've been very well able to put this into practice. Again, it is harder if use ICD-9 codes. We have our own internal system that we're able to pull it out.

As far as your question with oligo and increase in stillbirths, I know there was

a recent article in ACOG and we have a rebuttal to that that will be submitted.

We looked at our data last week, looking at stillbirths and we have not seen an increase since we have put this in. So just so everyone is aware of that and that will be out, if it gets accepted.

CO-CHAIR RILEY: Can we vote?

DR. WINKLER: Also, Kate Chenok is on the phone. Kate, did you hear the discussion and do want to enter a vote for importance for the elective delivery part of 39 weeks?

MS. CHENOK: I would vote, yes, it is important, thank you.

DR. GREGORY: Okay. We're going to move to scientific acceptability. And again, we had high on reliability, although two people said it was moderate. And then we were split on validity, four and four.

And, again, I think this comes back to the coding. One of the other issues

that we didn't address is that another exclusion is active labor or spontaneous rupture. And active labor in particular is not something that can very easily be detected through coding, although there are some algorithms for it.

But ultimately at some point after you get your denominator there is some chart abstractions that have to happen if you want to be as accurate as possible. And so that's where the issues come.

But, as a whole, we felt that it was relatively reliable -- that is was very reliable and relatively valid.

DR. BAILLIT: Can I just ask, in my hospital too we're having the same experience. We're going through by hand for all of these and after we screen them by the ICD-9s. Is it possible to build an exclusion for a classical section for those who want to do that?

Because we know that classical

sections are not equally distributed among hospitals. They're much more concentrated in the tertiary care centers. So is there a way to say, if you choose to go through by hand and pick those out, that we will accept that?

MS. MILTON: As a matter of fact, we can create a debt element that would look at prior classical cesarean and prior myomectomy. So that could be accomplished by chart abstraction.

DR. BAILIT: I would throw cholestasis in there, too.

MS. MILTON: Actually, the code has changed for that now. It used to be just liver disorder, now it's and liver and biliary disorder. So that has been addressed through ICD-9 codes.

Also there's two new codes out, if the coders are using it, that if the mother is in spontaneous labor with a planned cesarean section and they code for that then they no longer need to look at the record to see if

they were in active labor prior to the cesarean section.

So this is something that we're looking to build into our algorithm as the first check point. And if they don't code for that we'll still have the original checkpoints for cesarean section and then whether they were in labor or had SROM prior to the cesarean.

DR. BERNS: Yes, just -- I support this measure. I mean, I think this is a very important measure. But just to echo something that Kim said -- and Elliott, I think I know you have a response to this. But the denominator, the low risk women in collecting that data, again, is challenging.

And we're actually -- we did a pilot with 25 hospitals in partnership across five states and it's very difficult to do, certainly for all deliveries. And I understand that, you know, you can get a subset and that's I think the answer here.

But there are hospitals that do want to, you know, input data and monitor not just a handful of deliveries. So, you know, I completely support this measure, particularly the denominator, and having that be a low risk denominator, as opposed to, let's say, all early term deliveries or all scheduled inductions and c-sections. It is a challenge, for sure, I just wanted to make that statement.

DR. WATSON: I'd just like to ask a question. Again, I think this is a wonderful measure, and I think it's been so powerful that the country has changed its practice over a short period of time. We've seen dramatic improvement, and so I applaud the developers.

My issue in trying to educate physicians about this is, again, is the glaring exclusion of the previous classical and the previous myomectomy.

And so did I understand you to say

that we don't have to wait till ICD-10? That's there's a work around? And what was that work around and is that going to be something soon?

MS. MILTON: You would have to review the medical record. It would be chart abstracted data element that we could create and build into the algorithm as an exclusion to the denominator population.

DR. MAIN: Perhaps if this group recommended that, it might be considered more strongly.

I think as we've tried to rule this out over the last several years in California statewide, we learned first of all that there's no such thing as low hanging fruit. This is hard. But it does change practice dramatically. The biggest difficulty here is clearly in data collection.

But I think what's happened is, as this was recognized and supported by ACOG and others, is that we've started to develop the

work around's that make it easier. To change the ICD-9 codes. To add the classical c-sections.

Actually, classical c-sections -- every hospital I talk to says that's an issue. But the actual number of prior classical c-sections is quite small. Even in a tertiary center it is only a handful a year, when you come right down to it.

With the other work, we're getting around it in many parts of California by being able to combine vital records and the ICD-9 codes, which gives you gestational age imparities, so if that collection is done then there's only about five percent of charts that need to be looked at by hand to look at active labor or ruptured membranes. If we change those codes then it can be done even more easily.

There is a problem with sampling, though, Scott, in that the denominator gets pretty small pretty quickly. If you have a

small denominator in a few cases you'll get wide fluctuations, month to month or quarter to quarter. So I have not been recommending that we do a sampling, because we've seen big, big fluctuations in our rates.

You know, as to whether the denominator is all 37/38 weeks, versus just the low risk members of 37/38 weeks. It's harder to collect the denominator that way, but it makes more sense if you're really only looking at the cases from the low risk to begin with. And I think that's why it got set up that way to begin with.

DR. BERNS: I know we're not on usability yet, but I think it certainly gets to an issue there, as well. Again, I support the measure, but it is a challenge for some.

DR. GEE: Would you both be able to speak to -- certainly there will be fallouts with this measure, because of the number of exclusions, patient population, et cetera.

One of the things that Elliot and I have talked about previously, is that five percent would be the fallout rate that would be acceptable. So you would expect that you would have some of these that would come up in the data.

Can you speak to -- and when you're doing quality improvements, sometimes that's a way to get around the prior classical or your prior myomectomy, is just accepting that if you're doing QI or payment reform based on this that, there will be a fallout.

Can you speak to that in your process in choosing that number and how that's worked in California?

MS. MILTON: Actually, we do not set benchmarks. We have not set a benchmark, but for those hospitals that do report this measure to the joint commission we set target ranges.

Target ranges are based on all hospitals from the previous quarter, how they

performed with this measure. So you have a mean and you have a standard deviation above and below. And that particular hospital's rate would be plotted on that graph, so they can see how they are performing relative to other organizations that report the measure.

This will vary from quarter to quarter, depending on how well the nation has performed. So this is actually a moving target range that we have set.

And the second quarter of this year we were around 14 percent, for all hospitals that reported.

DR. MAIN: The five percent came from Leapfrog, that also uses the same measure. In a big sampling in California hospitals, and part of them in a March of Dimes collaborative program, most of the hospitals have been able to get down into the under eight and many under five, to two percent. And that's been the experience in Intermountain Health. I think you're running

about 2.2 percent -- 1.9.

CO-CHAIR RILEY: Okay. In the interest of time I think we should move on here, and vote for scientific acceptability of this measure.

DR. BAILLIT: This with acceptance of the c-sections, prior classicals taken out or no?

CO-CHAIR RILEY: With the recommendation.

DR. GREGORY: We're measuring on specifications as they stand.

DR. WINKLER: Kate, do you want to vote on scientific acceptability of this measure?

MS. CHENOK: Yes, I think it meets acceptability. And if I get cut off, I'm fully in support of this measure.

DR. WINKLER: Thank you. Okay, 25 yes, one no.

DR. GREGORY: We're now going to talk about usability and feasability, and,

again, it was rated highly usable, with two people scoring moderate. And very feasible, with one person scoring moderate, everyone else scoring high.

There were some questions about the risk adjustment. I thought, if I recall the discussion, many people felt that it maybe should not be risk adjusted. And that especially the issue with race, sort of, detracted from that, from the ability to see disparities.

CO-CHAIR RILEY: Comments? None, okay. So voting on usability.

DR. WINKLER: Kate, did you want to vote on usability? High, moderate, low, insufficient?

MS. CHENOK: Moderate.

DR. WINKLER: Thank you. Nine high, 15 moderate, one low.

CO-CHAIR RILEY: Moving on.

DR. GREGORY: We talked about feasability. We felt that it's feasible. The

big issues we've already talked about, in terms of parity, gestational age, labor. It's a little intensive but it's feasible.

And so the overall rating was pretty much mixed between high and moderate.

CO-CHAIR RILEY: So can we vote?

DR. WINKLER: Kate, your vote on feasability?

MS. CHENOK: Moderate.

DR. WINKLER: Thank you. Three high, 21 moderate, one low.

CO-CHAIR RILEY: And then overall suitability for endorsement?

DR. WINKLER: Kate, your overall?

MS. CHENOK: Yes.

DR. WINKLER: Thank you. Twenty-five yes, zero noes.

DR. WATSON: Can we formally make the recommendation to include previous classical c-section and previous myomectomy? Would that be appropriate at this time to make that recommendation along with this?

DR. WINKLER: Is everyone in agreement with that recommendation?

(Chorus of yeses.)

CO-CHAIR RILEY: And so we are moving on to cesarean rate for low risk first birth women.

DR. WATSON: Okay. This is my metric, my measure. And I was actually quite glad to see that the low-risk first birth woman c-section rate has been changed.

I never quite understood that metric. I think the NTSV is much clearer. Obviously, getting our c-section rate down is something that has plagued us for -- I think the data said 20 years. It's been longer than that.

I remember my first month in private practice my division chief came up to me and said I needed to get my c-section rate down. And I said I have not done a c-section yet. And he said it doesn't matter, you need to get that rate down.

(Laughter.)

And obviously we have been unsuccessful, the rate continues to climb. And so any help that we can get in getting that rate down I think would be greatly appreciated.

Our workgroup thought that this was a very important measure, had high impact. And had high opportunity for improvement, and was felt to be very, very important.

ACOG mentions that they feel that this measure is probably the optimal measure for cesarean section rate. More so than the total c-section rate or the primary c-section rate. And I think previously that's where people have been trying to put their emphasis.

Early data shows that, currently, nationwide we're at about 27.7 percent. And I think that the rationale for this is that this measure really kind of looks at the process of labor or the management of labor. And that's where we really need to put our

focus.

So I'll stop there and ask the developers if they have any comments?

DR. MAIN: Just one comment, in that this is a measure that really focuses on the group of -- the sub-population cesarean sections has seen the biggest rise in the also 15 years, besides repeat c-sections, but has been -- labor c-sections has been the biggest contributor to the rise of c-sections.

And it's also the biggest source of variation among hospitals and among providers. So this is where most of the variation lies, is in this population of women.

DR. GEE: I just want to say that this is eminently usable, because it's looking at the thing that we can modify most easily, I mean, without getting into the VBAC argument and complicated argument.

It's simple to measure. We're putting it on our report card. It's useful,

I think, for both public reporting as well as provider reporting. It's easy to do, you can abstract it from vital records, which is a godsend. In terms of hospitals wanting to work with you, our hospital association is very supportive of us collecting this measure, supporting this measure because it does not put a large burden on our providers. I think it's a fantastic measure.

MS. PARTRIDGE: I would just add that this is one of the core measures recommended by Secretary Sebelius for use in the Medicaid program. For those of you who maybe don't realize it, the Medicaid programs collectively pay for at least half, probably now, clearly half, given the recession, of the babies born in this country. In other words, about two million a year.

And in the stakeholder group that developed those recommendations for the Secretary, there was wide representation of a variety of perspectives, including states that

were concerned about data collection and burden and so on. And it was overwhelmingly endorsed. So it is certainly highly recommended by the Medicaid programs.

DR. GELZER: This measure was first endorsed in 2008, and maybe this is just a silly question, but why -- it hasn't had impact on the rate. Or has it decreased the rate of increase? Can we make it required?

DR. MAIN: I think there's a difference between endorsement and then being used by, either insurance companies by linking the payment, by people like the Joint Commission that have picked it up, and other groups. So it takes a few years to actually get into play, is what we're seeing.

But I think you need an additional incentive to go with the measure to make change sometimes. I think this is one that's going to need some incentives involved.

DR. ARMSTRONG: And for what it's worth, we have about 200,000 deliveries a year

and the rate is stable. So there's been no increase in it. Stable at a very high level.

DR. CALLAGHAN: Same with the national data.

DR. GREGORY: I think this is one where we not only need system issues, but we need a lot of patient education about this measure.

DR. BAILIT: I think the nice thing about this, and along with the one we just looked at, the 39 week, is that it's all tied in with inductions. So there is something that hospitals know that they can look at if, they're found to be an outlier on this, there's an obvious sort of place to look to change practices.

DR. ARMSTRONG: We see huge regional variation in this rate, around the country. Lowest in the north central region. Very high in the southwest and the southeast. So there's something also about the culture of how we talk to patients, how physicians train.

Maybe that has to be addressed as well.

CO-CHAIR RILEY: And how patients accept what we say. So that's the other thing, is that I don't think there's some variation in what patients' expectations are. You know, some demand what they demand and others are willing to listen to reason.

CO-CHAIR RILEY: These regional patterns match hysterectomy patterns around the country so it's all tied in.

DR. WINKLER: Who has the handheld mic? Guys, don't you have it back there?

DR. PRATT: Hi, I'm Steve Pratt, I'm an anesthesiologist from Boston, and I would offer one caution. Epidural's got blamed for c-sections for a very long time.

When we all looked, when you all looked at, what are the things that we know that increase the risk for cesarean delivery. Control for all these things and then said see all that's left is the epidural, it must be that.

As it turns out it, probably means that pain was a marker for dysfunctional labor and those woman ask for an epidural and that's what the prospective randomized trials have shown.

The concern that I would have is that it means there's something that happens in labor that we don't measure very well. And can't understand very well and all those things that we try to say are risk factors for labor aren't the whole picture. And I'm just wondering how you're going to deal with that?

CO-CHAIR RILEY: Okay. Any other comments? Are we able to vote on at least the first part of this measure? Importance to measure and report.

DR. WINKLER: Kate, are you there?

MS. CHENOK: I'm here. It's important.

DR. WINKLER: Great, okay, thanks. Twenty-five yes, zero no's.

CO-CHAIR RILEY: Okay. Moving on

to reliability.

DR. WATSON: For reliability, our workgroup felt this was a pretty clean measure. The numerator is c-sections. The denominator is basically defined by the measure name with just a few exclusions.

There was some concern during our discussions about the stratification of the data by age. And if I heard correctly, during our workgroup conversation, there was some data to support that. And I would like to ask whoever in the workgroup had that data.

DR. DENK: Actually, I brought it up. A mild question about the risk adjustment by age. And Elliot answered me and then I went back looked at New Jersey data and he was, I see exactly the same thing that he does and I'm going to start risk adjusting as soon as I get home, for age.

DR. MAIN: There's a very dramatic, impressive straight line for nulliparous c-sections, and actually c-

sections as a whole from age 18 to 40. It's not something that begins at 30, 35 or 40. It's straight line correlation coefficient of about .96. It's pretty dramatically a straight line.

So that women at 30 have seven or eight percent higher rate than women at 24. Women at 24 have higher rates than at 18, and it's all the way up the line.

So it wasn't just women over 35 or 40 which is in California. Which is in San Francisco, my issue, but it's a straight line all the way from 18.

DR. BAILIT: Elliot is that true for people in spontaneous labor or is that really a reflection of the practices that the older women have more complications, et cetera, et cetera?

DR. MAIN: Well, we worried about older, about it being a self full filling prophesy about how you care for women over 35 or 40. That's why it was so dramatic to look

at women 18, 22, 24, 30. And it was a straight line in that time period as well, even up through the early 30's and 40's.

So we felt it was more biological with some issues, you know, there's self full filling prophesies and how we manage people who are older or more obese or so forth that that's hard to tease out in this.

But when you look at the earlier ages and see the same thing happening, that's when we thought it was real.

DR. ARMSTRONG: We did something similar and then took the infertility singleton patients which sort of our marker of the premium births kind of thing. And they just have higher rates that go up also with age. So there's a bump that the premium baby gets. But age, you see the same thing.

And did you also look at obesity, stratify by that?

DR. MAIN: We've done -- obesity gets blamed for a lot in this area. But the

big bump in obesity comes with morbid obesity. But again, that was very hard to tease out from what the physicians' intentions are from what is biologically related.

They weren't necessarily that much bigger babies. So we have not yet done a correction, or a risk adjustment for obesity. The one caveat I have would be morbid obesity may be would be a different one.

And that is increasing, but it is still a small proportion of all the obesity and overweight that are seen in the U.S.

DR. DRYE: I'd just add that in our outcomes measures morbid obesity is showing up as important but not obesity in general. Probably because obesity is so common at this point.

CO-CHAIR RILEY: Okay. If there's no further comment can we vote on the reliability and validity of this measure? Kate do you want to --

MS. CHENOK: Yes, I support.

CO-CHAIR RILEY: Thank you.

DR. WINKLER: Twenty-five yes, zero no's.

CO-CHAIR RILEY: Moving on to usability.

DR. WATSON: Well, the usability for our workgroup that we had had some discussion about the risk the adjustment that we already talked about. But with that exception we found it to be highly usable.

And I think to answer some of the previous questions of why haven't we seen improvement previously? And I think that from a user's standpoint, I think it was poorly understood at first. The low risk, first birth. I think was a bit confusing. And of course, we didn't know where the data was.

Now we've had some time to collect that and so hopefully the usability with quality improvement going forward will be demonstrated.

MS. BRANDENBURG: One comment I'd like to make is we have been collecting the 39 weeks measure for about a year now for the March of Dimes. And one thing we have seen is we have dropped our c-section rate from about 39 percent to 32 percent.

And the only measure we've done is the 39 weeks measure. So I'm wondering if once the 39 weeks measure goes into place if it won't affect this measure as well with the c-section measure. Because we have seen that.

CO-CHAIR RILEY: Okay. So I think we can vote on usability of this measure.

MS. CHENOK: Yes, it's usable.

CO-CHAIR RILEY: It's a high, moderate, low, insufficient information.

MS. CHENOK: High.

CO-CHAIR RILEY: Okay. Awesome, thanks.

(Off microphone comments)

DR. WINKLER: It's 23 high, two moderate.

CO-CHAIR RILEY: And moving on to feasibility, are there any other comments about feasibility?

DR. DENK: Yes, I just want to echo what Rebecca said before. From the point of view of a state official or somebody outside the health care system. This is exactly what I'm talking about, a measure that's harmonized with what we can do.

I can do this, I can't do the 39, I can't do it in the way that it's endorsed here. So, you know, I kind of wonder why one measure has so many exclusions and the other one doesn't. But that's for future thought. But thank you.

DR. MAIN: We specifically designed this measure to have two specifications that are really pretty much the same. One that's able to be used with vital records and one that's done in a typical joint commission manner. So it has a double set of specifications you'll see in the application.

DR. ARMSTRONG: So this measure has it's first births, right? So can you get first births out of vital statistics data?

FEMALE PARTICIPANT: Yes, but you can't get it out of claims data.

DR. GEE: Just to agree with Chuck, we ought to be thinking more along the lines from a public health standpoint in population, improvements in population health.

Of how to use vital records more effectively for measure creation, and whether if we do need to modify them. How do we do that regionally or nationally and create technical assistance to states who are tying to do this. Because it is very challenging. But this a perfect example of how that works very well.

DR. PROFIT: I guess what my question would be is what the correlation is between the two specifications that you propose? And to some degree if the correlation is really very high I wonder why

we'd need both.

If one of them can be done post hoc without anybody having to go in an extract data?

DR. MAIN: Not every state is there yet, in terms of why not. A lot of states haven't been able to release vital stats data or use vital stats data for these kinds of purposes. For many months to a year or two after the fact. And so that degrades it's value as quality metric.

There is some quality improvement work that may need to be done on some state birth certificates for presentation. In California we had to work on that some. So there is some work that you have to do as a state. If you're going to take this on a state to make certain your data is as clean as possible.

In terms of head to head trials we've not done a big assessment of that, but it is a lot more time consuming to go through

the charts to do this. Which again, is why it hasn't had wide acceptability in many states yet.

Because A, the joint commission is not required. And B, you know, unless you use vital stats you have to do chart review to do this measure. And that's the difficulty.

DR. PROFIT: To some degree it feels like we're shifting the burden, maybe this should be a statewide burden but we're shifting it to the hospital. I guess, I'm not sure if it should be a statewide burden or a hospital burden.

But maybe that's a worthwhile discussion to have. Because I guess what we're doing here then is shifting it squarely onto the hospital, right?

Well, the state if they do the vital statistics. But I guess the other states, you could kind of say like well, if they're not ready like go get ready.

CO-CHAIR RILEY: Okay. So if we

can vote on feasibility.

DR. WINKLER: Kate, is it high, moderate or low for you?

MS. CHENOK: Moderate.

DR. WINKLER: Thanks. Sixteen high, nine moderate, zero low.

CO-CHAIR RILEY: Moving on. Overall suitability for endorsement for this measure? Kate, yes, no.

MS. CHENOK: Yes.

CO-CHAIR RILEY: Okay.

DR. WINKLER: Twenty five yes, zero no's.

CO-CHAIR RILEY: Okay. Next up is appropriate use of antenatal steroids. This is another joint commission measure. And this measure, just to remind people was. the numerator is a full course of antenatal steroids prior to delivery. In that episode of care, I should say.

And the denominator would be babies between 24 and zero sevenths weeks up

to 32 and zero sevenths weeks. And in terms of, this is sort of a no-brainer actually.

In terms of the data, there's lots of it as well as clinic guidance from NIH, from ACOG, from everyone you can imagine. Suggest that the appropriate standard of care for these babies is to get a dose of -- or get a course, I should say. Of the antenatal steroids if they fall between those gestational ages with excellent evidence that it will decrease the risk of RDS and NEC for these babies.

So in terms of importance to measure and report, everyone in the group felt that it was high impact. In terms of improvement, or whether or not there was a gap in sort of what people are doing is that there was clearly room for improvement based on the data that was presented.

And although there's been some improvement, which was good to see, to your point Scott, there was clear data that it had

been tested and in some places it's seen some improvement. They're still a ways to go in this population.

And the workgroup also felt that it obviously met importance. I don't know if there was anything I left off that occurred on the workgroup in terms of this. I can't remember, honestly. Any comments or concerns? Jennifer.

DR. BAILIT: Can JCo comment on why it's a full course as opposed to a partial course? One of the things we see is, we see patients show up in late labor, we give them a dose of steroids but you can't hold them off for the full 48 hours.

MS. MILTON: I know originally when the measure was endorsed they were looking at just the initiation of it. But when our technical advisory panel met to review the measures they felt that we should be looking at full course.

One thing that we have done in our

measure specifications, is that we would look at patients, if they'd received, let's say, one dose and then they deliver before the 24 hour period that would be an implied reason why they didn't get the full course. So therefore they would be removed from the measure.

So that's forthcoming in specifications beginning with January 1, 2011 discharges. Because we did hear that loud and clear from the field.

DR. BAILIT: Okay. So just to clarify, starting in January of next year that if you deliver before the full course, that excuses you.

MS. MILTON: Excludes.

DR. BAILIT: Excludes you, thank you, from the measure?

MS. MILTON: That's correct. '12, I'm sorry, got about a month to go.

DR. ARMSTRONG: So it excludes you or it's a hit?

MS. MILTON: I'm sorry?

DR. ARMSTRONG: It excludes you?

MS. MILTON: It excludes you from the measure. They would be taken out of the measure.

CO-CHAIR RILEY: So just for the non-clinicians in the room, I should have mentioned this, I didn't. Full course means that you get your first shot, and then 24 hours later you get your second and then hopefully deliver 48 hours after the first shot.

So it is not inconceivable that women will come in, they'll get the first shot. While you're waiting around to give the second one they deliver.

So this is just for the non clinicians in the room. I should have mentioned that. And the other issue is that if you use dexamethasone then it's Q 6 hours and you're getting four doses.

DR. ARMSTRONG: It would just seem

that's appropriate care, so that you wouldn't be excluded, you would count as a hit.

MS. MILTON: You get taken out of the denominator.

CO-CHAIR RILEY: Right, so the point is we are saying that you shouldn't get penalized for something that you can't control. Because tocolytics don't work, so someone's going to deliver before your 48 hours.

DR. WINKLER: Question, Celeste, does the specifications that you submitted to us include that change?

MS. MILTON: Yes, it's in the data element reason for not administering antenatal steroids. It's spelled out in there, in the notes for abstraction.

DR. GREGORY: Can you clarify, I noticed that you went down to 32 weeks instead of 34 weeks?

MS. MILTON: Right, when we worked with this measure with the original measure

developer, I think that there's some evidence that they haven't had, I might get this wrong, premature rupture. That you could go up to 34.

So it wasn't as clear cut if you were allowing if and or. So the decision was made to simplify it and just look at that discreet 24 to 32 week period. That was the reason.

CO-CHAIR RILEY: So I think, wasn't it to sort of mirror the NIH consensus so that we wouldn't penalize people. Because the NIH consensus, that was a lot of the discussion was 32 to 34 and back and forth.

And no one could decide and think they arbitrarily said okay, 32 weeks and so I think this just sort of mirrors that.

DR. GREGORY: Well, the NIH said 24 to 32 for ruptured membranes, and then, I'm sorry, 24 to 34 intact. And said, question mark, ruptured 32 to 34.

So I'm just concerned that by

endorsing this measure then there's this 32 to 34 week group with intact membranes that -- I want to make sure we're not sending the wrong message.

DR. JALEEL: I'm Jaleel, from Parkland Hospital. And over there the constant held debate with the obstetricians because of some of the controversies that you have mentioned. One is hypertension and the other one is diabetes.

And the previous old literature which this was based on had clinical trials which were done in hypertensive women where there was an increase in fetal deaths, in those hypertensive women.

And also there are multiple other reasons why it might not be suitable for mothers who have diabetes. So those are two exclusions that our observations in our hospital have for the antenatal steroids.

Do you want to comment on that?

MS. MILTON: This will be looking

at the evidence and I'm trying to remember if I saw anything in the evidence that said that there were any harms. I don't believe in the evidence that we reviewed that we saw that.

So I guess I'm going to bring my expert up here.

DR. ROSS: Can you repeat that again?

DR. JALEEL: So in our hospital two of the exclusion criteria for antenatal steroids. One is hypertension and a second one is diabetes. And so the reason that you have given is that in the previous Cochrane analysis which was done they had not specifically looked into hypertensive women.

And one of the trials which had looked into hypertensive women, there was an increase in fetal death. And so there is a concern for fetal death from that.

And then diabetes, one of the concerns is yes, it can cause hyperglycemia, it may inhibit surfactant action and all those

concerns. So I don't know what the recent literature is, recent evidence is on that.

DR. ROSS: Yes, I think the Cochrane had that response. Most of the patients who are receiving this are on constant monitoring. And so there's not thought to be a very big concern about fetal death and the glucocorticoids. I think that's diminished in the literature.

As far as diabetics, certainly the glucocorticoids can put that diabetic in poor control, and that has to be dealt with. Never the less the benefits are thought to far outweigh the risks in that regard.

CO-CHAIR RILEY: So, Mike, do you have a response in terms of this, what Kim's bringing up terms of the 32 to 34 weeks gestation. And sort of where does that leave us with the intact membranes, I'm sorry, I should clarify.

DR. ROSS: You know, I think the reason we chose that, again, was to avoid

trying to break it out into whether the membranes ruptured, and when did they rupture. I haven't seen any trend that we're pushing patients or pushing docs not to give the steroids between 32 to 34 weeks.

So I would continue it at this 32 weeks for the simplification. I would perhaps add one note for the future. And that is the concern that we're overusing steroids. Which is happening across the country. And people are using it like water.

Every patient now in so many community hospitals as well as the academic hospitals are getting steroids the moment they walk in with a contraction at 25 weeks. And the vast, vast majority of those are not delivering within the seven to 14 days.

I don't know, do we have any, we noted that it's a seven day efficacy, but is that counted in the?

MS. MILTON: We don't look at that.

DR. ROSS: We don't look at that. That might be something to long term consider is the number of courses of the steroids as well as whether you're giving it within an appropriate time window of perhaps 14 days. But for future discussion.

DR. PROFIT: So I'm a little bit humored I think. In 2008, we a lot of discussion about this measure. And a very similar measure from Vermont Oxford.

And the main reason we ended up endorsing this measure was that it included the 32 to 34 weeks exclusion for women with ruptured membranes, because the obstetricians felt very, very strongly about this.

And, you know, on the other hand we tried to make the arguments like, okay, what's that we heard yesterday about 80 percent of VLBW infants already captured by the VON and they're already collecting this data.

But it was felt to be the more

scientifically pure and in line with NIH guidelines to address the 32 to 34 week-ers. So we had this big trade off between trying to be perfect versus good enough.

And the perfect one out, and now we're going back to good enough. Which, I'm not against that but I would just kind of caution about extra data collection again for something that like 80 percent of the VLBW patients already submitting somewhere.

I guess the other point that I would make. And I understand we're in line with the guidelines at 24 weeks, I can tell you for a neonatalogist how frustrating it is to receive a 23 or 24 weeker from the community that has not gotten steroids because it didn't hit 24 weeks.

And I guess my last comment would be that I feel like the evidence for antenatal steroids on outcomes on babies is probably stronger than anything else we have in neonatalogy. So, you know, I'm fully

supportive of having some kind of antenatal steroid measure.

CO-CHAIR RILEY: So are you suggesting -- what exactly are you suggesting?

DR. PROFIT: So I'm not suggesting, I know there's no Vermont Oxford measure on here to harmonize and to think about. It is probably because they didn't want to resubmit after the last discussion we had.

But I do think that we need to be cautious, and I think for this round, you know, I feel supportive of this measure. But I feel like if there is another round in the future where we would have, if you know, if the VON would resubmit.

I'm just trying to be very conscientious about hospitals not duplicating work. And I think that's just wasted effort. All of us who work hard on quality improvement and trying to make things better are competing for resources and hospitals to do anything on

the front lines.

And so if money flows into a duplication of efforts, you know, on measures that are very similar, you know, that money is not available to do good things for our patients. So I just want to be conscientious about that.

I don't think we can do anything about the 24 week-ers for this measure and that's fine, but I would maybe this could be brought to ACOG, I'm not sure.

But if you have representatives from ACOG who could, you know, like leave it the fact at what point it would be brought up to extend, you know, to measure criteria.

Or maybe to investigate whether this is going to be beneficial to babies that are just below 24 weeks.

CO-CHAIR RILEY: So I think though to speak to that last thing I think that we would be looking for some clinical guidance? Before you could measure whether or not people

are doing it?

I think we need some clinical guidance that's very clear, with compelling evidence, that says we should be ratcheting down what we're doing and, because all you're going to measure if you go to 23 and half or whatever the number is. Is a huge variation in opinion from neonatalogy as well as obstetrics.

And that's going to be a mish mush of who knows what it is. So I think if that's an issue then we need to take that back to both AAP and ACOG to take a look at that 23 to 24 week gestation and come up with some guidelines that all of us can live with.

DR. PROFIT: Yes, I'm just suggesting that maybe we can make a recommendation as a committee that that should be in an important thing to be looked at in the future.

Because it's almost like when we say, you know, when we enshrine it at 24 to

zero, you know, with this measure that, you know, it will drive practice even more so.

And so I feel like at least we can make a recommendations maybe that this is something that should be investigated studied or discussed at least. I'm not sure we're able to get a trial on these babies together.

DR. WINKLER: I just want to make a comment, this measure is up for it's maintenance review and as Dr. Profit mentioned, there has been a change. And so I think that's something you need to factor in.

Because in the course of the prior endorsement it had, you know, one set of specifications that now have been modified as you've commented on to a significant degree.

So that is an important aspect of its, you know, sort of maintenance review is how the measure evolves and changes.

CO-CHAIR RILEY: So can we vote on importance to measure and report?

DR. WINKLER: Kate, are you still

with us?

DR. WINKLER: Twenty four yes, zero no's.

CO-CHAIR RILEY: Okay. So moving on to scientific acceptability, I think within the workgroup we sort of talked about it already. We felt that the reliability of the measure was high. Five felt that and one moderate.

And then in terms of the validity, again, the vast majority of the work group felt that it was highly valid. I don't know that there was any other discussion, do people feel there are other things that need to be clarified? Feels that we can vote? Yes.

DR. DENK: There's a little corner discussion going on over here. The issue of excluding from the measure. Anybody who starts a course but doesn't have a chance to complete it before the baby is delivered. You know, that means that you don't get credit for starting a course, right?

And so while it's not a denominator either, I mean, you don't get credit for a lot of attempts at good health care. So I'm just wondering if you thought about the pros and cons of giving credit. I know it's defined as full course. And that's obviously a trade off. Do you want to comment?

DR. ROSS: There's both ends of that spectrum, one may be that the patient comes in, the doc does a great job and can only give a course, a half a course a few hours before delivery.

The other is you wouldn't want to just sort of allow that and say, "Oh, we forgot to give this." So an hour before she's about to deliver let's go give her a dose. So I think excluding them is probably the wisest for the present time.

DR. BERNS: So I kind of smiled at that and I need some help from my OB colleagues, you know the literature better.

But is it clear that a full course, two doses is absolutely necessary?

My understanding is there's a little bit of, that's not crystal clear in the literature, that one dose even 48 hours and beyond is, it can be beneficial. So I am concerned. You know Chuck's point is an important point. Plus, you know, this is locking us into a full course, pretty much, of two doses, right? I mean, if I'm reading this correctly.

DR. ROSS: The literature, as most of the OB studies is controversial. The vast majority of studies that were originally done through the report on the efficacy used complete dosages.

And yet there are select studies, that you are correct, show that independent of the full dose or even the time to completion may not have a difference of effect as compared to the full dose. On certain outcome parameters, particularly bronchopulmonary

dysplasia.

Nevertheless, I think the majority of the literature and obstetricians would believe that the full dose is what's really most effective for the prevention of at least the acute outcome of respiratory distress syndrome.

CO-CHAIR RILEY: Also, I think that we're at this point in practice where, I don't know how you're actually going to get the answer to your question.

I mean, I think it's a daily debate on every unit. When you get that person who comes in and you're like, she's not going to make the 24 hours, should we do Q 12, is there data to support that? I don't think anybody knows, some people do it other people don't.

And I feel as though you can see studies which, you know, they are small, you know, so I don't know that we're going to be able to answer that question.

DR. WATSON: I'd like to ask our perinatologists if you all, I mean you may have just answered with no data. But at national meetings this question has been posed.

What's the latest time that you can give steroids and the answer was just before the ears come out. So just, nobody really knows, and is it advantageous to give it at that late date? So I just don't know.

DR. ARMSTRONG: Do you have a sense in looking at the data for babies that are appropriate for steroids but get only one does. How often is it a quality issue versus just a timing, no opportunity.

MS. MILTON: We really haven't looked at that so I guess I couldn't answer that. It seems though that when we do get inquiries into the measure the majority of the time, it's because they have delivered prior to. You know, before they could get the second dose.

That appears to be the issue for the most part I would say. A couple of the cases where there was a documented congenital anomaly incompatible with life. So that would have been the reason why they wouldn't have needed the steroids as well.

DR. ARMSTRONG: Right, so by excluding them you're leaving good behavior. You're not accounting for some good behavior that hospitals could get credit for.

CO-CHAIR RILEY: Other questions?

DR. ARMSTRONG: To answer your question, I don't think anybody knows how short a time. Kim's whispering in my ear, she thinks she's heard four hours. I have no idea, I haven't heard anything.

I just give it until I see, literally the ears come out. I think the reason that you get that answer is because we cannot predict when someone's going to deliver.

I think that that's the main

issue. Is that we don't have good predictors. Fifty percent of the people won't, who come in with contractions won't deliver, but I don't know which 50 percent necessarily.

In which case that's how I think we creep into that. Everybody gets it the minute they have one contractions which I agree with you is going to be, you know, in a few years is going to be a real issue.

I'm not so sure that we can get at that though. People getting too much steroids, I think that we can get at people getting too many repeated courses. The same patient, getting repeated courses which we now have data to suggests that that's a bad idea.

But I don't know how you would ever be able to measure that people are giving it with every contraction for every woman.

DR. ROSS: I guess there's thinking in the future of the number of patients who received it and then did not deliver pre-term, or the number of patients

who have received one dose and then didn't deliver at least for two weeks later.

And that's where we're sliding into this overuse. Which has effects potentially on fetal growth and neonatal outcomes. Adverse effects.

CO-CHAIR RILEY: So if there are no further questions or comments. Can we vote on the reliability and the validity of the measure? Kate, are you back on?

MS. CHENOK: I'm back, I'm going through security.

CO-CHAIR RILEY: We're on scientific acceptability of the antenatal steroids measure.

MS. CHENOK: Right, it's one or high, whatever the highest category is.

CO-CHAIR RILEY: That's how you leave your license as you're going through.

DR. WINKLER: Twenty four yes, one no.

CO-CHAIR RILEY: And then moving

on to usability. I think in the workgroup we talked about, we talked a lot about the augmentation. But most people felt it was highly usable and were sort of split on feasability between high and moderate.

But for the most part we felt that it was feasible. I mean, I do think that the issue of chart review, and resources et cetera, definitely came up. There's no easy way of doing this currently. Lee.

MS. PARTRIDGE: I would just volunteer that a process similar to that used for developing the child and maternity care related measurements set for Medicaid programs took place over this past year with respect to a set of measures for adults who are enrolled in state Medicaid programs.

And this was one of the measures that actually ended up being on the final list that was sent to the secretary for her consideration. And ultimate, we hope, approval. But that process is still underway.

CO-CHAIR RILEY: Kate, this is a high, moderate, low, insufficient information.

DR. WINKLER: Sixteen high, eight moderate, zero low.

CO-CHAIR RILEY: And feasability. I already actually discussed it. But feasability can we vote on that?

DR. WINKLER: Six high, 16 moderate, two low.

CO-CHAIR RILEY: Okay. So finally. Are there any other questions, comments before we go on to the final?

Okay. So overall suitability for endorsement for this measure. Kate, are you back with us? She probably forgot to take something off and so the poor thing is now being searched.

DR. WINKLER: Push one more time, everybody, please. Twenty four yes, zero no.

CO-CHAIR RILEY: We are moving on to the final one of the morning. Healthcare-Associated Bloodstream Infections

in Newborns. Scott, you're on.

DR. BERNS: Thanks, Laura. Okay. This is Healthcare-Associated Bloodstream Infections in Newborns. We discussed three related measures yesterday.

This one assesses the number of staphylococcal and gram negative septicemias or bacteremias in high risk newborns. We went through some of this data yesterday but for some of those that weren't here, in terms of impact.

This is a significant problem, particularly for high risk infants. Very low birth weight infants in particular. And infections result in increased length of stay as well as increased hospital costs and charges.

And other risk factors include central venous catheter use prolonged time using parental nutrition. A prolonged time with mechanical ventilation.

This is also very well written. I

want to add as well. There's a nice summary here about effect preventive measures and quite a bit of detail on that in this write up.

In terms of opportunity for improvement, there are a couple of notes in here about how this rate varies widely across different centers in terms of health care associating infection rates.

They note between six percent and 33 percent and that educational interventions in particular are noted here in terms of being able to decrease catheter related blood stream infections from, well, very significantly, as noted here.

So in terms of the workgroup we petty much rated this in terms of importance to measure with four high, with one being moderate. We note that this measure is very similar to Measure 478, and so we do have this general agreement on the high impact.

And the issue did come up again

here in terms of when patients are transferred from another hospital that it was what we want to comment on. You'll also see, I think you brought that up in our discussion, if I remember correctly.

You don't have to comment on it, you can if you want to.

In terms of evidence we rated in terms of quantity, quality, consistency. Moderate and high. So I'll just pause there and see if there are any questions or comments for the developers.

DR. DRYE: Sorry, can you just remind us how you handled transfers?

MS. MILTON: For the larger birth babies we're looking at those that are, I'm blanking out. It's greater than fifteen hundred grams, that's it.

We would take anybody that was transferred in within two days of birth. They would be considered high risk. Because that typically would mean that the baby has

problems.

They're being moved from a smaller facility to a tertiary facility so therefore they would be part of the measure that we would evaluate them just on the fact that they were transferred in.

MR. GILLIAM: I wanted to ask, in the brief description it talks about staphylococcal and gram negative septicemia and then under numerator it talks about different ICD-9 codes not being familiar with those. Do you, you would also include enterococcal and fungal isolates as well?

MS. MILTON: I don't have it memorized but I know some of those are in the list, yes.

DR. DRYE: Sorry, just to go back to the transfers for a minute. My brain slowing down after all these measures. Can you tell what the POA code was there was an infection, I think that I remember that you can. In which case that makes perfect sense

to me, what you're doing.

MS. MILTON: Whether it was present on admission?

DR. DRYE: Right.

MS. MILTON: Right, yes. There is a way of identifying that through ICD-9 codes and we're working through that right now. Taking a look at how this could be facilitated.

DR. DRYE: Okay. I mean, in general what we've tried to do in our hospital based outcomes measures is we usually attribute the outcome to the first admitting hospital, which also takes responsibility for the management of the patient going forward.

An in this case I think it's tricky if the patient's coming from another hospital to attribute the infection to the second hospital unless you can be sure it wasn't present on admission.

MS. MILTON: For a hospital receiving a baby that's infected it would be

the principle diagnosis code so that's how you would exclude them from that. In a hospital that has the baby delivered there the principle code will be a V30 code. That occasions the reason for being admitted is that they were born.

That's always going to be the first code. Then we'd be looking at other diagnosis codes that would be present on admission to exclude them if they were born with an infection at the reporting hospital.

DR. DRYE: And you probably just can't know for sure, right? If that baby's been at another hospital, where they acquired the infection. I don't have a problem with it because there's no perfect answer to it.

I did have a separate comment, one of the things I thought it would be good for you to talk about is the variation in the measure.

So this is a risk adjusted outcomes measure and the challenge we have

with risk adjusted outcomes measures, where we go ahead and quantify the uncertainty associated with the estimates is that if we are careful to quantify uncertainties sometimes we just can't see statistically significant variation. So in 2B5.3 which is on Page 20. You report, and I think this is really helpful.

The range of scores and there is a range in the 90 percentile score is 1.64 percent and the ten to zero, almost all the way through this 70 percent is zero percent but there aren't statistically significant results from the target range. I assume you're using the 95 percent confidence interval there, but I'm not sure it's stated.

So I just, I don't personally have a major issue with that for this measure but I just think it's important to think about the way the results look on the measure for other people on the committee.

And options you can pursue that

are reasonable to me, they're not necessarily to everyone. If you're using a really, really tight confidence interval like 95, which you'd want in a clinical trial you can back off of that and you potentially will see outliers.

I don't need to know personally, 95 percent certainty that something statistically different from a mean. If I'm thinking about a quality measure. So anyway there may be some ways that you can adjust the reporting strategies so that you see more variation even without adjusting you're modeling.

DR. BERNS: Anyone else?

CO-CHAIR RILEY: Okay. So it sounds like we're ready to vote.

DR. PROFIT: Can I ask one quick question? Does that include early ons you're including early ons and infections, right? There's no, the baby doesn't have to be three days old or 72 hours?

So you're just including those

because you're thought is that the incidence is so low that it doesn't change the measure overall?

MS. MILTON: We're looking at newborns up to 120 days of hospitalization.

DR. PROFIT: You know theoretically, if the number is very low and you think it's insignificant that makes sense. But quite clinically I find it a little hard to attribute, like an early onset infection to a hospital when that is something that has already present.

When, essentially when the baby is a fetus. So if an infection begins before birth it's hard to attribute the outcome to the hospital.

DR. ROSS: Yes, we discussed that with cholio and Group B strep and E. coli as the early onset infections. But I can't recall the pediatricians' issues on this.

MS. MILTON: I'm trying to remember how this went when our panel met to

discuss this. Because we did have neonatalogists addressing exactly what you said.

And they looked over the list of the codes that we're identifying here and they felt that it would identify those that were born with infections. And I don't have the list in front of me right now, but that was the consensus of the panel at that point in time.

DR. PROFIT: The denominator exclusions identified the babies that were born with infections?

MS. MILTON: Right, yes, it's one of the first ones for the ones that are born with it. There's a couple of tables that identify them, I think 11.11.2 or 11.10.2.

CO-CHAIR RILEY: Jochen, do you want to look at that before we vote?

DR. PROFIT: I don't have all the ICD-9 diagnosis codes on my head but I feel like it would be important for acceptability

within a community to make sure that that is the case.

Probably the overall rate of early onset sepsis is going to be very low. So I doubt if it will even make a difference. But I think just in terms of general face validity I think it's something I think that I'm sure would have come up in your expert panel.

So I don't know if there's an easy way to just look at it quickly, but I don't want to hold up the committee for this because I'm not sure how big the impact is. I couldn't believe that you would not notice or gotten feedback on that.

CO-CHAIR RILEY: Craig has some help for us.

MR. GILLIAM: When I look at denominator exclusions it does talk about length of stay less than two days. So that would catch most of the early onset Group B strep and also probably the E. colis. I assume that's what the pediatricians are

referring to.

(Off microphone comments.)

CO-CHAIR RILEY: Can you use your mic please?

DR. GEE: Yes, I'm just saying that was accurate. I think that the exclusions are length of stay less than two days. I think that's getting done. If there's an infection in the first two days that is not excluded. That's my understanding of this measure, is that correct?

MS. MILTON: Only if they are discharged within the first days are they excluded. That's correct.

DR. PROFIT: So it does say in 2A 1.9, patients with ICD-9 principle diagnosis code for sepsis are excluded. Is that what you are referring to?

MS. MILTON: Yes, those would be for the babies that are transferred in with that infection. Because that would be the reason for the admission.

Would be the principle code, that would be the reason for the admission.

DR. GREGORY: But they do have POA documentation now so, it should be coded as present on it.

DR. DRYE: I think you're saying the same thing. So any principle diagnosis code is a reason for admission it's just explicitly on there already on admission.

DR. BAILIT: But every newborn has the primary diagnosis as newborn.

DR. DRYE: But this is a transfer in, so the initial, the reason for admitting the baby to the hospital is sepsis. So then it has to be present on admission. If it's in the primary diagnosis field. Because by definition, in retrospect the reason for which the baby is admitted.

DR. GREGORY: So that's great if you're a transfer that doesn't address the issue if you're an inborn baby who has GBS.

DR. PROFIT: I'm sorry I couldn't

remember those quite, but it does say diagnosis codes for newborn septicemias are excluded. So newborns septicemia what usually covers early onset sepsis.

DR. ROSS: The newborn GBS or E. coli, was excluded but again, I can't remember which codes we use for that.

DR. PROFIT: I guess this goes into the reliability of the extraction and how the code is used. But newborn septicemia is for early onset disease.

And there is like acquired, there's a different code for acquired disease. But I guess that would be a question about reliability of the abstraction in that case.

CO-CHAIR RILEY: So are we all okay, no.

DR. JALEEL: It seems like it's important to know the differences, whether it is there or not. The measure title itself says health care of associated with bloodstream infection. So if we don't know

that it is difficult to work one way or the other.

CO-CHAIR RILEY: I think he's saying that he sees it.

DR. JALEEL: Neonatal septicemia is used for both early onset sepsis and later onset sepsis. So newborn sepsis can be either late or early. So we don't know that.

DR. DRYE: I don't know if you have access to the table that's referred to at that part of the, where the exclusion is. It's table 11.1 because it doesn't seem to be in our folder of appendixes. Oh, someone has it, great.

DR. YOUNG: So I'll read it out to the group that's here. Table 11.10, newborn septicemia or bacteremia is septicemia or sepsis of newborn. And bacteremia of newborn.

Table Number 11.11, is newborn bacteremia, that includes Group D strep or enterococcus, staphylococcus unspecified, staphylococcus aureus, other staphylococcus.

Friedlander's bacillus, there's also klebsiella pneumonia E. coli or pseudomonas. That's Table 11.11.

Then Table 11.12, is that included in this as well? Okay, sorry so Table 11.10 and 11.11.

(Off microphone comments.)

CO-CHAIR RILEY: Okay, so the question is, we have to rely on our neonatalogist to help us figure this out. Is that an answer to the question or no?

DR. JALEEL: We still don't know the answer but I would assume that the developers have gone through this and looked at this carefully because, as the title itself suggests, that it is health care associated infection.

So early onset newborn sepsis should be excluded from that. I would assume that they would have done it, but we don't know that for sure.

So I think it is, assuming that

they would have done it, it is probably okay to go ahead and vote.

CO-CHAIR RILEY: So assuming that that is an exclusion. Then we feel comfortable voting on this?

DR. JALEEL: Yes.

DR. ROSS: And we'll double check that.

CO-CHAIR RILEY: Okay, is everybody okay with that?

DR. KELLY: Does the top of page 16 help you? The developers.

MS. MILTON: That's the calculation algorithm. I'm actually looking at the table right now.

Okay, it's Table 11.10.2 would be the table when they come in with the principle diagnosis of sepsis from another hospital. And we're looking at streptococcal septicemia, staphylococcal septicemia not specified, MSRA, septicemia, staphylococcal septicemia, pneumococcal septicemia.

CO-CHAIR RILEY: We're not disagreeing with the bugs, we're disagreeing with. We recognize that transfers, we're all on board with that.

MS. MILTON: Okay.

CO-CHAIR RILEY: The question is if you're born in that hospital and you got infected in labor, and you now go to the NICU's, the concern is if there isn't a code that suggests that you were infected in utero or you got infected in the first 24 hours. It's not the NICU's fault, it's the uterus, or I don't know.

(Simultaneous speaking.)

MS. MILTON: So for those babies that are born there the first, the principle code is the V30, then we'd be looking at another diagnostic code that would be a sepsis code from Table 11.10, and there's two codes. Newborn bacteremia and newborn septicemia. Those would be the two codes we're looking at.

However we have become aware that

it goes up to the first 28 days of age, that you would code for that. So we're looking at the present on admission flagging of the ICD-9 code to identify that.

CO-CHAIR RILEY: Does that work? I don't know.

MS. MILTON: This is the discussion we had with AHRQ as far as harmonization of the measure.

DR. JALEEL: It kind of works but doesn't really. Because when the newborn comes in you don't know whether that baby is septic. So you have, probably a diagnosis of evaluation for sepsis, rather than neonatal septicemia at that point. So you don't know that at the beginning.

MS. MILTON: There's actually two ways it can be flagged, where it's known or unknown. And we'd only except if it's known. So it would be the way that it's documented in the medical record is my understanding of that.

DR. JALEEL: Yes, the documentation in the initial of 48 hours until we get the cultures back will be evaluation of sepsis and not neonatal sepsis.

DR. PROFIT: I guess one way around that would just to look at day of life three or 72 hours and just call it a late onset sepsis measure. I think you're out of the worry about the contamination then.

DR. JALEEL: For once the two neonatalogists agree.

CO-CHAIR RILEY: That means we've been here too long.

DR. KELLY: Okay, I think I've got it. I don't know for sure. But the bottom of Page 15, 11.10, it says the stop processing so I think that's the initial septicemia table. So does that answer the question, I think.

To me it says if it's initial septicemia and newborn septicemia you stop.

DR. DRYE: I think we're just asking how clearly is that defined? Does that

have a time limit on it? Like it's in the first two or three days, that code? Or if it's on the first 30 days then you potentially, babies get out of the measure who actually had a hospital acquired infection. If it's day 15, they get an infection and that's coded as one of those codes.

CO-CHAIR RILEY: Can you turn the mic on, or just like move up and just talk into that. Because we can't hear you back there. That's okay, scoot your chair up.

MS. WATT: Well, there are cords here. And I hate to lean over everybody, but I will.

I just want to make sure that I understand what the issue is. And so please refrain me if I'm incorrect. But if I'm understanding correctly the question, is are patients, are babies who are born already infected eliminated from this measure?

And the answer is, maybe. And it's not a fault of the measure, it's a fault

of the identification system of the coding system. There is no way to identify somebody using a code who was infected in utero.

We've done the best we could to eliminate those people, the babies, with the secondary diagnosis of septicemia or some sort of an infection.

But if you all can share with us a good reliable method of identifying those patients who are born infected. That would be a terrific help to us, because we can't do it now. As far as we know.

CO-CHAIR RILEY: So, Jaleel, can you repeat, I mean, I got the impression that you were saying that if you just started at day three of life, instead of day.

DR. MURI: It doesn't matter, an infected person is an infected person and there's no way to say okay, no easy way, to say the infection occurred on day three of life.

Using the codes, what you have is,

you've got the V30 code and you have that secondary infection code and it's very, very difficult to put time constraints on that. Well like, impossibly difficult.

DR. ROSS: Do you feel that delineating it by bacteria, you know, if you avoided the E. coli and GBS would that give you more of the hospital acquired infection?

DR. PROFIT: I think if you avoided the GBS that would probably help. But I think the instance is so low it's meaningless.

MS. WATT: It's very low isn't it?

DR. PROFIT: Yes, it's like one in a thousand.

DR. ROSS: And E. coli

DR. PROFIT: E. coli sepsis is less prevalent, so you'd want to capture that for sure.

DR. JALEEL: Listeria is the other one, but again, the infection rate is so low, it's mainly the E. coli.

But you would not be able to take the E. coli out because E. coli can be late onset infection as well. So I don't see an easy way out. That's different from the Vermont Oxford network measure which only looks at late onset. And I'm not sure how they do it.

DR. PROFIT: I think I would assume that overall the correlation between the measures would be very high. But, you know, I feel like that is maybe work that should be done.

MS. KIEHN: I guess the one other question that I have is have you done a crosswalk to ICD-10's?

MS. MILTON: Oh yes. We actually sent it in to NQF, they have it if you're wanting to look that over.

DR. WINKLER: It should be in your files.

MS. MILTON: Right, the infection codes, there really wasn't a lot of

difference. But boy, the surgical codes, there's 42,000 now.

DR. PROFIT: I think just a general, because our effect is for every infection measured to is transfers out, so it's like the exposure time the baby has in the hospital. I think that's going to introduce bias on all of the measures that we have before us.

I don't know an easy way around that. There probably isn't. But, you know, it's just something to consider as we go along here.

CO-CHAIR RILEY: So I guess the question here is, giving it back to you, Dr. Profit, I mean, this is going to be the perfect versus good enough. Is sort of what we're talking about.

I mean, and if the neonatal infection read in the first couple of days from the uterus, or wherever it comes from, is really, really small, are you still going to

get meaningful data out of this measure, is really what we're talking about.

DR. PROFIT: I think I would want to say yes, but I'm not like, I think my confidence is not terrible high about it. And I think there will be substantial, there might be at least. Substantial concern with the neonatal community about this.

And I think you'll hear a lot of feedback potentially about that. So, you know, I think it's a valid concern that people have.

I'm not sure if you actually tested it how whether you'd truly find a big difference. But I think that will be one of those areas that people will try to, or will at least note as a concern about the measure.

DR. JALEEL: If you look at, I'm not an epidemiologist so I don't know what the incidence of early onset sepsis is. But if you look at he patients who are in the NICU, many of those extremely low birth weight

babies will receive antibiotics for five to seven days.

For either proven sepsis or clinical sepsis. Suspected clinical sepsis. So there are many babies who do get that and if you're coding that, that will be coded as neonatal sepsis.

CO-CHAIR RILEY: So maybe what we should do is go ahead and vote, at least on the first part of this and then I think some of these concerns about how valid it is or how reliable will come up? Is that fair? So let's vote on the importance to measure and report.

DR. WINKLER: Kate, did you rejoin us? Twenty yes, four no.

DR. BERNS: Okay, should we keep going?

CO-CHAIR RILEY: Let's forge ahead.

DR. BERNS: Great, we're going to forge ahead here. In terms of scientific

acceptability, we've talked a bit already about the numerator and denominator.

But in terms of harmonization and comparing to 478, the exclusions are a little bit broader. Specifically around a length of stay over 120 days, and being enrolled in a clinical trial. That's my understanding unless something has happened since I last read this.

And then in terms of the group, before I get to that. Just getting to reliability, so further liability studies are bing done, I guess now. And I thought it was interesting, of 26 contracted measurement system vendors. I didn't realize that there were that many.

CO-CHAIR RILEY: There are more than that.

DR. BERNS: There are more, okay. And the other thing that I noted in the validity section here is that a couple of the exclusions are not addressed in the

literature. But were included to harmonize with other measures from I guess, CMS and perhaps others.

And those were specifically length of stay under 120, length of stay over two, sorry, the other way. Length of stay under two, length of stay over 120, and being enrolled in clinical trials.

In terms of the group, and we talked about some of this already. In terms of the transfer piece and I guess in the denominator statement. Experienced death, there was some concern about hospitals that have worse outcomes potentially looking better in the measure.

So overall in terms of reliability we were mostly in the moderate range as a group. And our review of validity also on the moderate range. Questions, comments?

DR. DRYE: I just have a general comment. As a participant, as a measure developer and other context. Putting measures

through NQF. So validity, I think NQF sort of tightened up or been more explicit by what it means by validity.

And in this very well written application, the validity that's presented is that JC takes feedback from users on a regular basis about, sort of the accuracy of the data. And I think data, the codes used to capture what we're trying to capture. And that's important.

But NQF is requiring now that face validity, that expert panel assessment of measures be quantified systematically. So we take votes on our expert panels now and really ask, you know, do you think this measure is valid in the sense that it captures the underlying quality construct that we think it's trying to capture.

And it's interesting listening to the dialogue here. I'm not sure that sort of ad hoc feedback from hospitals is going to get at that, right? We're hearing concerns from

neonatalogists here that maybe this is capturing, you know, what it really needs to be capturing.

And I just wonder if any of you can comment. Because I think these forms are not the most current NQF forms and they don't say if you're using face validity you need to quantify that. Those are rally new forms that we've been using.

DR. WINKLER: These are the new forms, these are them.

DR. DRYE: So maybe you just don't see. You see it when you're filling it out, but it doesn't print out those criteria, okay. But, I mean, do you have any thoughts on that because I think we're kind of loosely using the term validity in this committee.

DR. WINKLER: Yes, it's an evolving concept because, you're right. The attempt is to try and address validity, we try not to be overly prescriptive about what that means. Because there are many ways to assess

validity.

Certainly face validity done in an organized systematic fashion but construct validity and certain kinds of, you know, test retest kinds of analysis assess validity.

So, you know, we're looking for something more than, yes, it looks good to me. And so but it's an evolving thing. And you're absolutely right, I think there is a lot of our membership in the audience who really want to see things, you know, a little bit more objective, and a little bit more quantified, to the degree that that's possible in this world.

DR. DRYE: And I think another option listed on the NQF form now is comparing outcomes with other measures. And we're going to be talking about this set of newborn infection measures.

And, you know, I think Dr. Profit made a suggestion. Maybe they should be arraigned against each other and see whether

they really truly are capturing similar things. Because we're struggling with all of them a little bit because they're so complex.

DR. WINKLER: That's your agenda item for after lunch.

FEMALE PARTICIPANT: If we ever get there, I couldn't resist.

(Laughter.)

CO-CHAIR RILEY: Okay. Are we ready to now vote on reliability and validity of this measure?

DR. WINKLER: Can everybody push again, are we missing anybody? Thirteen yes, 11 no.

CO-CHAIR RILEY: Usability? Do you want to talk about that, Scott?

DR. BERNS: Okay, in terms of usability. Let's see, there was a note in here about consistent improvement in aggregate performance measures. I guess between quarter two in 2010 and quarter one in 2011.

There was some data presented in

here. And in terms of feasability, let's see in terms of the group, just to go back here. Pretty much we were in the moderate to high range in usability and feasability split between high and moderate. And a note about needing to harmonize with 478.

MR. GILLIAM: If I could speak to the usability. And since we're talking about public reporting. The issues about public reporting are not using necessarily this measure, or even VON, they're using NHSN. Which is slightly different and is more specific and a narrower focus. And as these efforts to decrease these events occur then the numerator becomes more and more important.

And even though Group B strep or maybe early onset E. coli maybe uncommon, if you're talking about smaller and smaller number of cases then using the codes. And I accept what you say, that you can't differentiate between true health care associated and what's reported.

That may be more of an issue, and I think neonatalogists and infectious disease people would have more of a concern about this from a public reporting standpoint.

DR. PROFIT: Have you compared the results of these measures to ones that are based on like, blood cultures, you know, where you're, because I just wonder how good the coding is, you know, in relation to like some sort of like former evidence on blood cultures.

MS. MILTON: We haven't done any comparisons to any other measures on this particular measure.

DR. JALEEL: I have to reiterate that the number of babies with less than 1500 gram, who get antibiotics in the first five to seven days for sepsis is a significant number.

DR. PROFIT: I'd just want to thank the non clinic folks who give some perspective. I feel like there's a lot of charting that goes on that's highly variable.

You know, a baby in a NICU may, you know, have some minor symptoms, be started on antibiotics, a new physician comes on. Doesn't know the baby that well, say well, let's just continue the antibiotics.

Like I don't know this baby so, you know, not that that's necessarily good quality but, you know, cultures nay have all of been negative. And so you know there's different measures, like different measures like the CDC has different criteria, and the VON and your diagnosis.

So there is some uncertainty about what truly constitutes and infection or not. And I think once it goes to public reporting everybody is going to asking like, well what's this? Are these truly infected babies?

Or can we train our coders only to be very specific about they're actually going to abstract. Or are we going to train our doctors to only specifically write down, you know, X, Y and Z.

So I'm just wondering whether some, you know, whether if it's only based on coding whether there is going to be a lot of gaming there in the system about, you know, the rates are going to look better but are we truly going to make a difference to the patient?

DR. BERNS: Okay. Any other comments?

(Off microphone comments.)

MS. BRANDENBURG: I have to agree with im as well. I mean, if you're basing it on coding, I would be concerned about coding alone, because that's not always been proven the most accurate.

DR. BERNS: Anyone else?

CO-CHAIR RILEY: So can we vote on the usability of this measure? High, moderate, low, insufficient?

DR. WINKLER: One high, ten moderate, 13 low.

CO-CHAIR RILEY: Moving on to

feasability, is there anything more we need to say about that? Can we vote?

DR. WINKLER: Push again just for good measure. one high, ten moderate, 13 low.

CO-CHAIR RILEY: Okay, moving on to the overall suitability for endorsement.

DR. DENK: Sorry, quick question, this is harmonized with another measure we did yesterday? Did we approve that one?

DR. PROFIT: Yes.

CO-CHAIR RILEY: Okay, so now we're voting again.

DR. WINKLER: Can everyone just press one more time? Okay so that's where we are on the first go around. And remember we've talked about this. We've got multiple similar measures that's our afternoon agenda item. Is we're going to put these side by side and take a look at them.

I think there's been an evolution in discussion and thinking and consideration. And this will give you an opportunity to

regroup on these measures, okay?

CO-CHAIR RILEY: So we're going to take a 15 minute break.

(Whereupon, the above-entitled matter went off the record at 10:58 a.m. and resumed at 11:18 a.m.)

CO-CHAIR RILEY: Okay, we've got we've got our work cut out for us. Dr. Gee, let's go.

DR. WINKLER: We're still having that echo.

(Off the record comments.)

DR. WINKLER: All right. The next little bit is going to be a little bit of a departure to what we've done. We're going to be looking at a composite measure.

And just some history. There is a growing interest in composite measures. Because they do kind of bring summary information together. Several years ago NQF had looked at the measure evaluation criteria as applied to composite measures.

Now admittedly, we haven't reviewed that in a couple of years and there have been a lot of new composites made. So there are some challenges that we have yet to confront.

But composite measures are typically measures that are some way combined of individual components to arrive at a single score. So that's what we mean by a composite measure.

And one of the sort of underlying principles is that the component measures should either be an endorsed measure or meet the criteria for endorsement. The measure we have in front of us is the measure 1769, an adverse outcome index.

In this particular case there's no indication or desire for the individual components to be endorsed independently. They are components of the overall composite measure.

However we do want to be sure that

the components are appropriate. That they are, you know, they contribute to the quality construct. They are appropriately defining coded or whatever the data collection requirement is.

That within the context of the composite they are usable and feasible. So it's going to be, so it requires you to think a little but differently.

Then the actual composite measure, the overall, there are some additional criteria to consider. Because the composite methodology has additional elements to it. So again, even if an individual component is not, you know, super strong on it's own, as part of a composite that may be perfectly fine.

The other issue is to think about the construct for quality of the entire composite. And is that well understood and do all the components, are they consistent and contribute to that quality construct, okay?

So that's what we're thinking of

when we're looking at the importance to measure report criteria around composites.

Scientific acceptability, you know, the specifications should be just as good but you will need all the additional specifications for standardizing scales, if necessary. Scoring rules, weighting rules, how you handle missing data of some of the components.

You know, sampling issues if necessary. So the criteria are not different but you have to look at different elements when you're looking at a composite measure, because it has different characteristics.

We're still looking for reliability and validity testing. We're still looking at the resulting score, demonstrating meaningful differences.

Ideally we'd like to see some analysis of how the components contribute to the overall score. And also that the scoring and weighting rules are consistent with that

conceptual construct.

So these are the kinds of elements that are additional evaluation criteria when we're talking about a composite. In terms of, I think usability and feasability, there's not a whole lot there.

The one key element is that the data is collected and maintained such that you can de-construct and go back to the component measures to figure out what were the contributions of various components to the ultimate score. So somewhat different.

This particular measure I'm going to ask Kathleen who's the lead to kind of describe the measure conceptually in general. And then what we're going to want to do is quickly go through each of the components to see how you feel about the contribution of that. Are there any issues?

` Thinking about all the usual criteria but within the context of being a component in a composite. And we'll vote,

just take a single vote on each of the component elements to see how everyone feels it is as a component.

And then we will wind up by talking about all of the different evaluation criteria for the composite measure. The usual importance, scientific acceptability, usability, with the additional criteria that the composite evaluation criteria has.

So, complicated. We have a lot of steps to go through. So with that I'm going to go.

CO-CHAIR RILEY: Can I ask a question first?

DR. WINKLER: Sure.

CO-CHAIR RILEY: As we sort of go through these, I recognize we're going to look one at a time. Is there the opportunity to say no, to a particular component, yet the whole rest of it looks okay? Or how does that?

DR. WINKLER: I think that's going

to become one of your issues, because in all honesty if you don't like one of the components and feel that that's the fatal flaw of the measure, and it's part of the whole, then it's going to be very hard to justify recommending the whole with a fatal flaw as one of the components.

So I think you do need to think about your evaluation of the component in the context of being a contributor to the overall composite score.

MS. PARTRIDGE: I was unclear, and I think maybe it would be helpful at the outset if this was clear for all of us. Whether we're being ask to evaluate only the adverse outcomes index or whether it's all three components. Because in the submission, sometimes I think they assume all three. The AOI, the weighted one, and the severity index.

DR. SIMPSON: I think that hopefully in the introduction that Kathleen's going to do and then perhaps with some

comments from the developers we can get that context before we launch into further discussions.

DR. SIMPSON: Well, actually I think that was a good question because in reviewing the material I did not have a solid answer to that question. So maybe you should mention that right away before I proceed.

Are you intending to, does the measure, the AOI, the weighted and the severity index, is that what we're discussing? It wasn't clear.

CO-CHAIR RILEY: Do you want to come to the table?

DR. SIMPSON: Yes, why don't you, that's very helpful to know because I did not know that and did not glean that from the materials.

DR. DRYE: And is that an all or none? Or if we're, is the committee considering each one? Because that's a little different than unraveling the index. It's

just saying, you know, the AOI versus the weighted for example. Or can we consider them separately?

DR. SIMPSON: I think you can discuss that, sure.

Okay, so the adverse outcome index, just as a background was created in the concept of a research study. And looking at the potential impact of team training in labor and delivery environments and their outcomes.

And this was a randomized trial, that was done almost ten years ago and in the context of this study. There were two consensus groups of physicians and nurses who came up with these ten indicators.

And then later on the indicators were weighted, from my understanding, through a ACOG quality improvement patient safety committee. So you had your two consensus panels and then it got weighted.

Now there are three then issues. So there's the adverse outcome index, the

weighted adverse outcome score, and the severity index. And they are all interrelated.

Now, they include these ten components. In hospital maternal death, and interpartum neonatal death, they all have, I'm not going to read into all them exclusions and clarifications.

Uterine rupture and unplanned maternal admission to the ICU. Birth trauma and it's noted that it's not the same birth trauma as the AHRQ measure, although we did not approve that yesterday.

And then unanticipated operative procedure, admission to the NICU of a baby over 2500 grams. Apgar score less than seven in five minutes. Maternal blood transfusion and third or fourth degree laceration.

And the lastly behind this is that most of things are very rare and so it you combine them together you could be able to use less of a sample and get an idea of what was

going on.

Now the philosophy is that each of these things individually represent a aspect of substandard care. And that was repeated throughout the application. I'm not sure if that is a true statement, if all these represent substandard care.

Some of them actually might represent quality care. So there's some disagreement about that. The developer also mentioned that there were several studies in which this was used.

They said three studies, totaling about 50,000 births. And then gave a list of references. I was able to pull all those references. And of course there's the original randomized trial of team training.

And then there was a study of the hospital, the Beth Israel Hospital that did not participate in the team training but coordinated the team training study. So that was, first I thought that was the same data

but it wasn't. So those are separate data.

Then there was another study that was done. Again a single hospital looking at the active management of risk at term. With the theory that if you induce a subset of people who are at risk for a potentially adverse outcome you might be in better shape.

And they used multiple things to evaluate their intervention and adverse outcome index was one of them.

Then there was the Yale study by Petker in which they looked at interventions over time. A comprehensive obstetrical team training program, patient safety program, multiple intervention. Fetal monitoring certification, induction at the appropriate time. The addition of a patient safety nurse.

I can't remember them all, and in full disclosure, I was the one that recommended they do all that. And I have to say that I was part of that. I did not participate in the study but I was one of the

consultants that went there and told them to do all that stuff. That they eventually did do.

So naturally I was thrilled when I saw that study and saw that things were moving in the right direction based on the AOI.

Then there were two additional papers published. These papers their specific intention was to evaluate the AOI, not to evaluate any intervention measuring using the AOI as a unit of measurement.

The one was Benedetti's group at the UW in Seattle, they looked at just the one hospital and they were going to see if the AOI was handy could it be used to for the entire state. So they spent quite a bit of time just looking at their own hospital's data and going back and forth looking at reliability and validity and comparing medical record review versus the algorithms put forth by the National Perinatal Information Center, is my understanding.

And so I actually emailed these folks to see what the decision was because they never said at the end of the article what they had decided to do.

And they said that they had decided not to go forward using the AOI statewide adoption. So that was additional information.

Then there was a recent paper published by Hamilton. And they were looking at how does the AOI perform? They used a four hospital system, with about 7,000 births and they were trying to figure out does the most common thing that happens really lead the, I guess, the trend of the AOI.

And their contention was that third and fourth degree lacerations the Apgar situation and maternal transfusion were leading indicators. And in fact the third and fourth degree lacerations being the most common might obscure some other things that were going on.

And also other people did find that third and fourth degree lacerations in admission to the NICU were the two most common things. And so they then might obscure other trends and other important things that are going on with the other eight or nine indicators.

There was also a couple of opinion articles published that had some concerns about use of the AOI. There was not anybody that said, wow, this is a great thing, we should do it. But they were measured in their, you know, their enthusiasm.

Then there was another article that was cited several times, however I could not get it. And you might be able to comment on that. It was something in Wyoming, and it was cited on the A1 website I could not find it.

But it was something about the hospitals in Wyoming. I did spend quite a bit of time trying to find it, so I don't know

what that was about.

So the main thing is that this is the philosophy here is that one or more of these things is a indicator of substandard care which is, that's controversial. And that together they can be handy because they don't require as big of a sample.

Now I understand that the national Perinatal Information Center has been using this with a lot of their member hospitals. And you did mention that several times here. I didn't get, I mean, none of that data is published. So I was unable to, you know, make a big comment on that. So you might be willing to say something about that.

So they include both process and outcome measures as well. So that's, Reva asked me to stop there and then we'll discuss further once we talk about these individual things.

So that's the state of the evidence at this point.

CO-CHAIR SAKALA: Thanks, Kathleen, that was a great summary. Would you like to clarify now, the relationship of the three different options and whether you're putting forth one of them specifically for us to consider?

DR. PRATT: Hi, I'm Steve Pratt, I'm an obstetric anesthesiologist. I was lead anesthesiologist on the original prospective randomized trial. I'm sort of here for our team. We have two or three other folks an the phone as well. So they'll be asked to help as we move along.

I actually think that all three are important and here's why. The three different components are, one the adverse outcome index, and that's the percentage of deliveries that is associated with one or more of the ten events within the component.

So if you have all ten it still counts as one on the numerator. The next, each one of those adverse events has a weight

associated to it.

Ranging from five points for a third or fourth degree laceration up to 750 points for a maternal death. The sum of everything below maternal death is equal actually to maternal death.

And that was one of the caveats we said in weighting this is the worst thing that can happen is to have a mom die. And even if all the rest of it occurs it couldn't be worse than a mom dying.

There are two ways to use those weighted scores. One is to take the sum of all of the points from all of the events that occurred over a period of time and divide it by the number of deliveries. That is the weighted advert outcome score and it gives you a sense of how much general bad things are happening on the unit.

The second is called the severity index, and you take that same numerator of points, the sum of all of those points from

all the events from all deliveries, and divide by the number of deliveries that had an event.

So in other words the numerator from the AOI. And that speaks to how badly the events are when they occur. And so an example of why this is important is one could get rid of third and fourth degree lacerations by cesarean delivery on everybody.

One would expect that the adverse outcome index would then go down. But lots of other bad things might happen and the severity index therefore should go up. Because the adverse events that would occur would exclude that bottom, the lowest scoring event.

And so we didn't want to allow there to be policies and protocols that might decrease the rate of events but worsen their severity.

So I really think that all three, and we've talked about this a length. That really all three measures should be considered here. And it's a way to measure quality in a

more aggregate way.

DR. SIMPSON: Can I ask a question about that? I was looking at the Yale study where they just counted the worst thing that could have happened. So people have operationalized it different ways.

What I couldn't understand is, let's say I was looking at my own maternal mortality data over the last ten years trying to figure out how this would relate.

So if you have somebody that maybe has four or five of these things but it's one case. Are all of those things then added into the, so you could have a couple of bad cases and really all of a sudden have a very bad score then, right?

Even if none of those cases related in any way fault. Like an MBA comes, well she's pregnant she gets a blood transfusion, her baby dies, she goes to the ICU. Now she dies.

And that's a lot of bad stuff

happening just for that one patient and it had nothing to do with your so called substandard care. That's a very concerning situation.

DR. PRATT: So again, it's the reason to have both of, all three of these. So you're right, the weighted scores would move quite heavily with a couple of those types of cases. Versus, now the adverse outcome index in fact wouldn't. Because that would count as one or two cases.

(Simultaneous speaking.)

DR. PRATT: One MVA would be one case. And so it's a way to in fact to deal with all of those issues. The Adverse Outcome Index wouldn't change with that bad case.

The other part of it this is, as you've already pointed out. Many of these bad events are in fact probably not preventable. And we know that, we accept that.

That's true of most adverse events in medicine, and if we look at them preventability is probably on the order of 50

percent for most of these terrible adverse events.

We don't have a great way to try to identify those that are preventable. But we understand that if we can work on that 50 percent that is, we can still have significant movement.

And the Yale study showed on the order from the beginning of their study to the end a 25 percent or better improvement. We showed similar results. And that's worth something.

DR. SIMPSON: Can you explain how the weighting was done, you know, who decided what weights per event, and how was that done?

DR. PRATT: Susan, are you on the phone? Is there any way to figure out if we have callers?

DR. MANN: Sorry, I was on mute. I'm Susan Mann, I'm also an obstetrician gynecologist at Beth Israel Deaconess Medical Center and one of the authors of the

participant and the original team tracking trial.

We ask the American College of OB/GYN patient safety, quality improvement patient safety committee. And I believe it was a consensus process, I don't sit on that committee so I wasn't privy to the results of their voting and the process they did, but I know it was also a consensus scale.

DR. PRATT: And like I said earlier, the only role we gave them was the sum of the first nine events could not be more than maternal death. After that they came up with the specific numbers.

DR. SIMPSON: Okay. And then in your consensus panel, your two consensus conferences, that's how you came up with the ten. Where there, it doesn't mention in all the articles, or at least maybe I skipped it. How many did you start with and, you know, what did you weed out to get to that ten?

DR. PRATT: Yes, so we actually

published that process a few years ago.

DR. SIMPSON: Yes, I got them all right here.

DR. PRATT: We started with more than fifty. And again, Susan, speak up because you were the lead author on this.

But we started with more than 50. We used six fairly specific criteria as we went through each one of these. Many of them overlapped. And the goal was, what we ultimately wanted were events that we thought, one were likely to be preventable.

To some extent we were looking at things that were likely to be preventable by broader aspects of care than a single policy. So things that would change the overall quality of care.

And so we were looking specifically at teamwork. But overall other protocols and broad guidelines have now been demonstrated to be effective, as you've mentioned the study at Penn.

Looking at management of risk in late gestation. But we were, at the time, looking largely for these that might.

DR. SIMPSON: For teamwork, right? They were developed to test teamwork.

DR. PRATT: That was originally the development.

DR. SIMPSON: Okay. That's what I thought.

DR. MANN: And also for the ten process measures.

DR. PRATT: Yes, those have not been included here though. These are just the ten outcome.

DR. SIMPSON: Does anybody else have any questions about the actual measure?

DR. PRATT: If you'd like I can, on the development you commented that some of these might be actually measures of good care. And the two that sort of jump out are maternal transfusion and admission to the intensive care unit.

We would never want to say, oh, it is a bad idea to transfuse a mom with, you know, with a hematocrit of 12, right? And we understood that as we developed it. The issue there is that many of the events on labor and delivery that lead to either one of those are associated with bad care.

And I'll use the ICU admission for, the leading causes of maternal ICU admission as you look from study to study, to study, are maternal hemorrhage, eclampsia, preeclampsia. Those tend to be the leading causes.

The literature is very clear that we don't manage any of those things very well. Both in looking at what we know about actual care but also when we look at the simulation data. These are things that none of us do very well. Or that are associated with a relatively high rate of inadequate or substandard care.

Rather than trying to have to go

through each one of these various reasons for which a mom might get admitted to the intensive care and measure each one of those. Which would have become much more difficult. We went with the concept that if these high risk deliveries are managed effectively the likelihood that mom will end up in the intensive care unit goes down.

And so this was a way to sort of grab all of that substandard care in these high risk deliveries and identify when they've gone badly and mom has ended up in the intensive care unit.

Similar thinking went into the went into the transfusion conversation. And, you know, management of induction, for instance. Would be a risk factor there potentially. And that was the thinking behind those two in particular.

DR. SIMPSON: Okay. Well I did look at the references you provided as far as ICU admissions and number one was hypertensive

disease. Then hemorrhage, then septic abortion, and non obstetric sepsis.

And so those again, in my opinion, are not all amenable to care. I mean, stuff happens. So if those are the leading causes of admission to the ICU for a mother and that is an indicator then that would be something controversial, I guess.

And then overall philosophy, is that if you look at some of these things there could be sort of gaming of the system if you were sophisticated enough to say, well hey maybe I'll just bring the people down to my L&D unit rather than admitting them to the ICU because I'll get a better score.

Or, you know, she's on the borderline, let's not transfuse her because, I mean I don't know if that would really happen. But it does seem to kind of push people towards doing things that might not be in the best interest of mothers and babies.

Also with the avoidance of the

third and fourth degree laceration maybe you're going to have more cesarean. You know this is just a philosophical thing when you look at the whole thing as a composite. So I wanted to bring that up as well.

DR. PRATT: That is fair, and you know, we worked on this very hard thinking about, you know, god forbid we would ever try to dissuade people from taking very sick moms and putting them in the intensive care unit.

We had such a hard time believing that anyone would actually do that. And, you know, you speak to, yes, bad things happen with severe preeclampsia and the hypertensive disorders.

On the other hand up to 80 percent of moms with severe preeclampsia do not have their blood pressure managed well. And you know, I think about my own specialty as an anesthesiologist, this is something in which we are absolute expert.

And my guess is that many of my

colleagues around the country are not as involved in the management of those patients as we might be. Again, now speaking to the teamwork discussion.

Is it possible that, you know, some moms are running around with an hematocrit of 18 and feeling a little bit fatigued and they didn't get transfused because of the adverse outcome index. Would that be possible? I suppose so it would. But I would certainly hope that that would not be big.

DR. SIMPSON: I'm not saying it's legitimate, I'm just throwing it out there as something that's in the literature, so I wanted to bring that up in fair discussion here. Others have mentioned that in the literature.

DR. MANN: Similarly, the NICU admissions. Boy I have a bad echo. Do you want to talk about the NICU admission commission?

DR. PRATT: Yes, this is a biggie, in some places it is almost a primary driver for their adverse outcome index and, you know, Janet can probably speak to this better than I.

When we look at the data for some of these hospitals there's some folks who have adverse outcome index rates of 20 or 25 percent. Now remember we are almost entirely driven by NICU admissions. Now remember this is at term, and greater than 2500 grams.

So these are good sized, term babies, and they have to be there for more than 24 hours. So this isn't they just went and got some blood cultures for rule out sepsis. These are generally real admissions at term.

And there's such huge variation in that that it's one of the, that's the one that I think is going to be also difficult to deal with. I frankly think that those hospitals should have to own up to the fact that they're

sending 25 percent of their babies to the NICU.

DR. SIMPSON: Don't you think that there could be some financial incentive for that?

DR. PRATT: Yes, I do, I absolutely do.

(Simultaneous speaking.)

DR. PRATT: I think that that's a good thing though to put that in the numerator though, that these folks are making those decisions for financial reasons.

DR. SIMPSON: It's possible, I don't know.

DR. PROFIT: So I had the privilege of training at the BIs so I know the system very well. But the hospitals, they're structured differently and some hospitals babies will be sent to the level two unit to receive two days of antibiotics. And in other places they'll just have a work up and they can go to the mother's room.

And so I wonder whether some of your high rates actually derive from these just different structural setups. You could argue that well maybe they should change the setup but I think, I wonder whether those are two differences in neonatal outcomes or just in the way the hospital is structured.

DR. MANN: It would still be a neonatal outcome if the mother and baby aren't bonding or breastfeeding because the baby's in the NICU.

CO-CHAIR SAKALA: So maybe we should save the discussion for the individual components as we go through the measure.

MS. PARTRIDGE: In your submission you mentioned that, and I'm not quite clear what the it is. I think it's probably all three measures. Are currently being used in QI collaborative's, both in the state of Maryland and the Premier System.

Now the Premier System we're all pretty familiar with I think and Maryland's

one of the fairly progressive states in improving maternal and child health quality. Are they using them in the context of reinforcing the team concept?

DR. MURI: Thank you Steve, I just wanted to introduce myself. I'm Janet Muri, president of the National Perinatal National Information Center.

The Maryland patient safety program used it in the context of team steps, IHI, bundles, NICHD. So they did some collection around those components as well as some process. And then they used the AOI as well.

That whole program has just finished and they're in the process of publishing. So they're working on some articles on that. And they found really nice improvement for many of the hospitals.

The hospitals that didn't improve, seemed to have some issues around the team, the leadership of the team and things like

that. Where they really couldn't really fully engage the team concept.

The Premier is a subset of their hospitals that are part of their excess risk program. They have a patient safety program that they've been involved with for about two or three years.

And they too are beginning to publish. They've just moved over to an AHRQ Grant to do some research with a comparison group of hospitals that are not doing the interventions that they're doing.

CO-CHAIR SAKALA: Jennifer.

DR. BAILIT: My concern with this measure in general, and I want to give you some examples and just see what you say, is the risk adjustment piece.

So for example if everybody had the same mix of patients at all these hospitals that would make sense to me that this would be a good measure.

But given that we have trauma

centers, and given that especially now there's a movement towards accreta centers where those women are, we know that they're a high risk. Some are going to die, they better get transfusions. They should be in the ICU's.

And so these patients are not equally distributed across America. I can understand using AOI to look at your own hospital year after year after year and see how you're doing. Because your case mix doesn't change that much from year to year.

I have grave concerns about comparing hospitals using this with no risk adjustment. Can you speak to that a little bit?

DR. PRATT: I'll again, let Janet speak to that because we've been able to this not risk adjusting at the individual patient level. But at the type of institutional level a little bit. And comparing similar institutions.

DR. MURI: Yes, one of the things

that we've looked at when we've looked at the Maryland patients safety data as well as the Premier data is to divide the hospitals by level of care. One, two and three, and then also by academic, non academic.

So that was, it's not really a risk adjusting, we've talked about really risk adjusting at the patient level and I think that's the direction we'd like to go in. But in terms of just looking at differences in types of hospitals we can subset the data and look at the metrics that way.

DR. BAILIT: So that makes sense to me to some degree to stratify and look at within the stratas. But all tertiary hospitals are not alike. A tertiary care hospital in the suburbs is different than in the inner city. Is different than the two or three accreta centers around the country.

So I just caution that we maybe under risk adjusting or under stratifying because you really need to know a lot about

those individual hospitals in a community before you can say one is worse than the other.

CO-CHAIR SAKALA: Nancy.

DR. LOWE: I have a question, really which gets to one of our key criteria, which is the public reporting aspect of this thing.

You know indices that are developed for use in research and are composite measures are usually fairly complicated for even the most sophisticated among us to interpret. Because of the complexity of the individual elements that go into a composite.

And that was my reaction to this whole thing. Is if we throw these out to the public and I think Jennifer has brought up some of the issues with it as a public reporting measure.

But I have great difficulty with understanding how, at the public level, this

could really be well understood. In the context of the complexity of not only patient mix but provider mix.

You know a rural hospital or a frontier hospital in the middle of Oregon, which has 49 babies a year. If they have one maternal death and it gets reported. Which may be a totally - or a non preventable maternal death, an MVA off the highway, a bad preeclampsic who is fulminates and comes in and it is a tragedy. It just, I am really struggling with the public reportability aspect of it. And the understandability from a public perspective of this complex of a measure.

So if you could just speak to that a little, why you think the public would be able to resonate with this measure and really understand it substantively. Rather than us approaching it from the individual measure standpoint.

DR. PRATT: For me, and, Susan

I'll let you go, and actually I think we should all probably comment. My response to that is I think it's going to be hard to figure out what the public was going to be able to understand.

I'm not sure that they can necessarily understand even the component measures. Within the context of the complexity of health care.

I think having some idea of stratification and we've done that with help, number one. Number two the, again, that single maternal death is not going to change their complication rate very much. It's going to change the severity score a lot.

But the complication rate is not going to change any more than the third or fourth degree laceration that they had last week as well.

And then finally while we get very concerned about those sorts of things the likelihood that there's going to be a maternal

death in a place who's delivering, who know, if the maternal mortality rate is 0.8 per 10,000.

The likelihood that if they're doing 200 deliveries a year, that's 50 years worth of deliveries. So the likelihood that we're going to run into that particular problem very much I think is quite small.

Yes, we should look at the multiple deviation, standard deviations out in the bell shaped curve of these events. But you're probably now talking in that particular example, you know, four or five, six standard deviations out in terms of that rate.

The question is do you want this to be something that works very well in the middle and a couple standard deviations out or does it need to work all the way out to the very ends of those statistics?

CO-CHAIR SAKALA: Chuck.

DR. MURI: I'm sorry, I don't know if I can really add too much to that except

that I appreciate that concern. And I think that, you know, I think there are a lot of measures out there that the public has difficulty understanding.

And so, but I don't think, I think part of the public is, you know, are payers and other quality groups that are doing assessments and reviews and things like that.

So I think if it's a measure that maybe that need a little bit, or measures that need a little bit more explanation I think you can do it in language that makes it clear for these three.

But will everybody understand it? Will the whole general public understand it? Probably not.

DR. SUTHERLAND: I guess I'd like to go back to Kathleen's point about the issue that we could bypass a lot of these parts of the composite by doing cesareans.

And I'm going to go through the list. Interpartum fetal death, uterine

rupture, birth trauma, unanticipated surgery. Apgar less than seven and third and fourth degree lacerations.

That's six out of your ten composites. And there really aren't any balancing things to look at the maternal morbidity that we're seeing from all the excess cesarean that we're doing.

A lot of us, I think, understand that some of the complications we're seeing in obstetrics are higher in the last 20 years because of our interventional attitudes.

And I know there are folks around the table that really support natural childbirth, so I guess I would like to bring that up to kind of think about.

What the individual things in the composite are and are they really balanced? As far as what our public health goals are for women's health. So I'd just like to hear some comments.

DR. PRATT: Susan, do you want to

take some of that?

DR. MANN: That's why we have a, I'm sorry, the echo is bad. That's why we have the weighting, is to look at that. And furthermore the experience at our own institution collecting this data, actually our cesarean section rate has decreased in the last several years.

So I think that it's not a necessarily, people don't make an individual decisions to do a cesarean section based on their rate of - on the average outcomes index score.

It's a very complex decision of why you do a cesarean section, so I am a proponent of natural childbirth as well and vaginal deliveries and I don't think that these aren't necessarily in concert with one another.

I still think it's information that been useful for hospitals. We certainly see hospitals drill down and do quality

improvement projects such as hemorrhage protocols, shoulder dystocia drills based on their experience with the staff.

DR. PRATT: Just one thing to clarify as well. I thought that we had made the language clear that the neonatal death is actually in hospital neonatal death, not interpartum.

And if it says that I'm sorry. I thought we had changed all that. It is not just interpartum, it is during the delivery hospitalization. So that changes that outcome significantly.

Because babies can die after cesarean delivery or due to events related to a cesarean delivery. But obviously if they have an elective cesarean delivery they cannot die of an interpartum death. So that one would be changed at least a little bit.

CO-CHAIR SAKALA: Thank you. Chuck.

MALE PARTICIPANT: Do I have time

for one additional comment?

CO-CHAIR SAKALA: Yes.

MALE PARTICIPANT: (telephonic interference, unintelligible)

DR. PRATT: The other part of this is, I think we're all in favor of having normal vaginal deliveries, and as an anesthesiologist I'm not willing to use natural childbirth. Although my wife had both of her babies with no medications, thank you very much.

But, on the other hand, if we can demonstrate that the cesareans that we're doing are in fact changing adverse events. If the complication rates are changing because of the cesarean deliveries, or at least in association with those, then those are the ones we want to be doing.

And I'm not convinced that doing cesarean delivery to save a third or fourth degree laceration is necessarily the right thing to do.

But if one is doing the c-sections and it's improving the rest of these scores, those are the ones we want to do. That is the kind of medicine, at least I think, that we likely want to do.

And as Susan said, our cesarean section, that rate went down at the same time that our adverse event rate went down. And our severity index went down. So it all moved in a good direction.

CO-CHAIR SAKALA: And are you attributing that to this program implementation in particular, or other things going on at the time?

DR. PRATT: We really believe that this was, that the teamwork was a big chunk of this. Now as Susan alluded the teamwork actually ended up leading to many, many, many, other changes. As you sit there and work as a team you realize all the other issues that come up on the unit.

And so the development of

communication protocols with the attending anaesthesiologist came out of this. So if there is a vacuum delivery now on our unit, I get called.

So if there is a stat section for that failed vacuum I am already in the operating room. Ninety percent of the time or more the vacuum works and I go back to bed at 3 0'clock in the morning.

Those are all things that are not specific to teamwork but came out of all of this. Hemorrhage protocol, shoulder dystocia protocols. Drills to help practice all of these things grew out of all that.

So I won't say just the team steps based teamwork training did it.

DR. BAILIT: I think there's a difference between, you know, if we measure health care and we put a lot of sunshine on it you have impetuous to get better and things are going to get better.

But I think the NQF has purview to

hold measures to a higher standard. Which is public reporting of inter hospital comparisons.

And that's a different kind, potentially a different kind of measure. Even if this is very good at what you're talking about quality improvement.

It doesn't necessarily mean it's ready for the big stage of we're going identify hospital X as being a not so good hospital and Aetna won't cover it and Medicaid won't favor it.

And they have big implications, and so that quality improvement work and something that's useful in an individual hospital quality improvement work is not the same as what we need to hold the measure to.

CO-CHAIR SAKALA: Chuck, please go ahead.

DR. DENK: Well I can think of a couple positive things to say about this. I agree with a lot of the comments about there

is a public reporting measure of hospital quality it sort of has a lack of transparency.

And part of that is, I think, because, you know, I think you're overdoing the link. I mean, we all want to have our indexes relate strongly to a single construct.

And your construct that you've named is quality care. But I don't think, I think this thing can have value without going that far. And it's sort of an intermediate thing.

What's striking me about this is that it solves some problems that I've been thinking about for the last two days. We have problems with sort of defining what are good validation strategies for a lot of the measures that we've been talking about. The process measures that we've been talking about today.

We have also been talking about the fact that a lot of the outcomes we'd like to prevent happen very rarely and don't

aggregate well to individual hospitals and things like that.

And so this is a sort of, I think of this as more of a utility measure. Where it's not necessarily for public reporting at all. And it's not necessarily labeled as a quality of care thing.

But the constructed measures I adversity, and that it can be used exactly in the way that it's been used in the literature described. You know, to assess the generalized impacts of multifaceted things.

Everything, you know, from improving teamwork in hospitals. In communication in hospitals to what's the impact of delayed prenatal care or low quality prenatal care.

Or you know a mom moving from one state to another in the middle of her pregnancy. I men, you know, there's a lot of issues like that. And so this could, something like this could become a good sort

of standard.

And I don't know if that means we would endorse it. But it could become a good standard for, you know, use this to validate generalized measures to prove a broad impact across a lot of outcomes.

And let me just close with another example. I'm sure a lot of you saw the same article I did a year or two ago. I think it was called an index of near miss, near miss mortality or something like that.

Where a whole bunch of different adverse events were added together and the scale was, or the weights were basically as far as I remember, how likely each event was likely to result in a mortality. I think it was all maternal mortality.

So you know, a laceration would have a very low weight because it's unlikely to by itself to lead to it, right? And you know a lot of things, and there are a lot of things in a near miss that aren't in here.

What's nice about this is that it, you know, it's kind of randomly poking at different parts of the system and so it would be a good first pass way to look at systemic improvements.

And it would be a good sort of generalized way as a first pass for validations, maybe. So those are comments I have.

I think this thing could have a place in the firmament of measures but maybe not, it doesn't have that linkage to direct patient quality improvement that we've been using as a standard for the last two days.

DR. WINKLER: Just to clarify, you know, public reporting is one type of accountability. And while that may not be the only it's certainly something of an area we see a lot of.

But other accountabilities around the payment accreditation, blah, blah, kind of things are also included. So it isn't

absolutely public reporting but that accountability realm.

DR. DRYE: When you say that, Reva, I just want to clarify, because I think this is really important. I totally agree with Jennifer's point.

And if you move into the payment realm I think you're in the realm of comparing hospitals in this case. And so that's really what we're asking.

Even if you don't use the word public reporting. You're saying should this be used to compare and either publicly report or quantitatively reward or penalize hospitals even with this measure. Would that be fair? So it has to meet that standard.

DR. ARMSTRONG: Yes, and I would say as the payer in room, or one of the payers in the room. That from an absolute count perspective here. The number of complications per delivery, you know, it's interesting, it says something.

I think when you put the weighted averages on it, you know, you really have to dive into whether that's the right weighting and then the risk adjustment piece is huge.

Because you can't compare hospitals, it's a non starter in a discussion about hospitals until you can risk adjust it.

CO-CHAIR SAKALA: Kathleen, did you?

DR. SIMPSON: I just wanted to make a comment that to respond to what you said about it is a measure of adversity. But the one thing I think that can't be overlooked is that there's an underlying assumption that this adversity is related to substandard care.

And that is not the case in each one of these things. If that were the case, if there were somehow, because I really like this concept, you know, I'm not opposed at all.

I like the whole thing, it's just that if there's some way to do a preventable

versus non preventable for each one of these ten things, and I know there's not. It would be great, but there isn't a way.

And again, the way this is put together is you're making an assumption that there is a direct link between one of these ten things and substandard care.

And it's repeated throughout all of the things that you've submitted. And I don't think that is true. You know, so that's the problem I have with that. The adversity is not all related to substandard care.

CO-CHAIR SAKALA: Kathleen, I just wanted to mention, one of those examples I think is the incidence of episiotomy versus tear. And so episiotomy being the direct clinical intervention where we're wanting to avoid third an fourth degree lacerations.

If you have a shoulder dystocia you need to get the kid out that's not a clinical mistake. Whereas episiotomy ought to be used rarely and with caution.

So why, how did you choose third and fourth degree lacs, and some of these things. What were your considerations in thinking towards that?

CO-CHAIR SAKALA: I think we need to save the specific questions for later. Because two of the three versions are so influenced by the weighting process, I wanted to ask Sean Currigan if you could comment on that? About how that worked.

MR. CURRIGAN: I was not present for the weighting because that happened before my time. But my understanding was that it was how measures were developed in those days. A bunch of guys around a table, and they discussed.

I think at that time, I mean it wasn't that long ago. Everybody, they discussed it and deliberated it over, like half of a committee meeting, about the weighting. And I don't know that they used any numbers to come up with those weighted

scores.

CO-CHAIR SAKALA: Okay. Thank you, that was helpful. So I think we're going now ask our lead discussants to very briefly present each of the components.

Can we try to stick to five minutes for presentation and quick discussion and a vote. And my understanding is that it would be possible to vote no on some subset and then revisit this at the end. Looking at the whole and say is it a deal breaker?

So it's not as if we vote no we stop. So we'll go through them all then we'll come back again to the total. So the first one is Bill Callaghan for in-hospital maternal deaths.

DR. CALLAGHAN: Hi, I think this will be brief, a lot of what has been discussed already pertains to this particular measure.

So the measure is all pregnant women who died during the hospital. They come

in for delivery and the denominator is all pregnant women who delivered during whatever time you're measuring this in.

As somebody who has spent the last ten years writing and talking around the country about pregnancy related mortality and severe morbidity. This is something that's near and dear to my heart.

I view maternal deaths as a true sentinel event that demands that the reasons for death, that the factors associated with the death be identified, reviewed and that that information be used to take action for preventability.

That being said, I share the same issue that has been brought up over and over again in terms of how that relates to accountability.

We view, we do surveillance of maternal mortality in the division of reproductive health. It's a little bit different from National Center for Health

Statistics.

We identify about 600 deaths per year, probably about 400 of them are during the delivery hospitalization. So these are extraordinarily rare.

And that being said, we think that at the same time that fact that one women died likely resulted from processes and system problems that occur over and over and over again all over the place. And in fact don't result in death.

Perhaps because of the sheer tenacity of young women who can survive a lot. But that doesn't mean that those processes shouldn't be addressed. So our take, my take, I guess is personal on maternal deaths.

These should be reviewed in aggregate. Perhaps at the facility but even better at state where these systematic errors, if there are systematic errors. And again, we think that about 50 percent are preventable by changes in health care behaviors, or system

problems.

And that those, the factors associated with them, once they are identified, those are shotgunned out to everybody. You don't assume that it's that hospital that made the mistake, and it only happened there. It happens everywhere all the time.

And the system needs to be improved over and over again everywhere. So from that standpoint, these are incredibly important events to review but I don't know that they should be, we should, as everybody, as many people have said, being used for accountability and public reporting.

I have some real problems with that, the review of maternal deaths being used in that way I think they should be reviewed but not just not counted and ticked off.

CO-CHAIR SAKALA: Okay. Thank you.

DR. CALLAGHAN: Counting and

ticking them off the information gets lost, there's nothing that can be learned from it except perhaps maybe in the one instance that it occurred. Perhaps a public health perspective as opposed to an individual facility perspective.

CO-CHAIR SAKALA: Okay, a quick comment please from the developer perspective, and then we'll hear from our other.

DR. PRATT: I would just say that these are not, we in no way would suggest that these are mutually exclusive. We are no way suggesting that this is the only way that maternal deaths should be reviewed.

And I agree with you that they are such high sentinel events absolutely every one needs to be reviewed. That doesn't mean that they couldn't be also part of a numerator process for adverse events.

So that would be my only comment back to you was that our intent was never to try to undermine the case review for this.

CO-CHAIR SAKALA: Comments from other members of the committee on maternal mortality as a component?

Okay. Wow, we made our five minutes, almost. So let's have a vote then about this particular component.

Okay, great. So next we move on to Rob, for uterine rupture

DR. WATSON: Well, my component's uterine rupture and I'm not sure exactly how I feel about this, so let me just sort of summerize.

It's a very rare event, it occurs in about 0.06 percent to 0.55 percent in the general population. It occurs in a greater frequency for VBAC's at about 1.5 percent.

The developers feel that uterine rupture could be associated with substandard care, being especially reflective of management decisions during labor, induction of labor or augmentation of labor.

And they feel that the rate of

rupture is an indirect measure of clinical decision making. ACOG rates this third out of the ten measures on their weighting list.

I had some concerns originally about the feasibility and maybe confusion of data collection between a true uterine rupture and a dehiscence, which we see quite frequently. But I think you all have worked that out.

So it's something that happens very, very rarely, you know, it can happen in spontaneous labors as well. Which would not be any kind of substandard care necessarily. And it could be prevented by a c-section.

CO-CHAIR SAKALA: Thank you. Comments on uterine rupture as a component of this measure?

CO-CHAIR RILEY: So I think this is one of those that, this could drive us exactly where we don't want to go. I mean, I feel like this is one measure where you put this out to the public, and not even the

public.

I'm thinking myself, okay, if I'm going to dinged for a uterine rupture because I allowed someone to labor because that's what she wanted and I had everything in my institution to support that.

I'm not going, I mean, it's not going to be that long before I get smart and say you know what there's no benefit to this hospital to do VBAC's and we'll be done.

So if we want to close that door, which I understand there's a lot of people that want to close that door and make that access. Then I really think that putting this out will get us, we're getting there anyway. But I really feel strongly that this will probably close the door on it.

DR. DENK: Just really briefly to that I agree with the logic of that but when he said it I thought, okay the cesarean you want to prevent is the primary one. And that way you don't have uterine ruptures. I mean,

that's a good way to prevent uterine ruptures, is not to do the primary.

CO-CHAIR RILEY: That makes great sense, but the reality is that, and we're working on that, we have a measure. But the reality is that this will also take away any patient who feels that she wants to give it a go. She's not going to have any access.

CO-CHAIR SAKALA: And 20 percent of child-bearing -- of pregnant women now have a history of cesarean.

DR. WATSON: I don't think, furthermore, from a quality standpoint, this is a rare event, and each one of these would be scrutinized at the quality committee at the local hospital. So it's not like these things are going to happen and no one's ever going to pay any attention.

They'll be looked at and if there has been misuse of Pitocin or mismanagement of labor that would be addressed at that local level.

CO-CHAIR SAKALA: Other comments about uterine rupture as a component?

DR. GEE: Would it be appropriate to speak to why would it be before maternal admission to the ICU, it seems like uterine rupture is weighted too much.

Can we speak to the weighting, it seems like it's too heavily weighted. Because it's somebody's, and often it's not avoidable. If you want to let someone do a VBAC you may have a uterine rupture.

And we discuss that with patients. That's somewhat predictable as an outcome with VBAC. So I think it's weighted too heavily.

CO-CHAIR SAKALA: So we could go on record as saying that, that we can't negotiate to change the terms of what we're considering. Are we ready to vote, 37 seconds, you're good. Vote.

So on that we have four yes, and 20 no. And I think we didn't report to the people on the phone the first measure was

evenly divided, 12 and 12.

So now we're going to go to Elizabeth for unplanned maternal ICU admission.

DR. DRYE: Okay. So this isn't my area of expertise, but I'm going to both this one and the NICU admissions. And I do develop outcomes measures for a living. And I just wanted to in 30 seconds comment on this as a outcomes measure going through NQF, it isn't risk adjusted, we've talked about why.

If you think 50 percent are preventable and 50 percent are not it's really hard to understand why the measure is not risk adjusted. There's a requirement at NQF, if you're not going to risk adjust an outcome that you show that risk adjustment is not needed.

And there is no data in the application that I saw anywhere where risk adjustment, for example, for patient factors like age even were tested and dismissed.

So that's a huge thing and another thing is that usually for outcomes measures we try to characterize the amount of uncertainty around our estimates because outcomes are affected by quality and by chance and by patient factors and we like to try to at least capture the, quantify the amount of variation that's probably due to chance so that we are comparing providers fairly and not by chance calling some providers worse than others. So that's just as a preface.

And then finally when we're moving from a QI measure, and it sounds like there's a lot of expense with this as a constructive QM measure towards, a measure for comparing, and then we have to pay attention to how the outcomes are defined and whether those outcomes are coded, as we talked about a lot in this group I think in the last two days, systematically across institutions.

So I'm going to be speaking to that as I look at this measure and then the

NICU admission measure.

And so briefly this component is unplanned maternal admission to the ICU. Which has a frequency as reported here of 0.1 to 0.4 of deliveries.

And the rational is that the causes are typically preventable including hemorrhage sepsis, hypertension, pulmonary edema. The numerator, I think one thing Reva encouraged me to look at since I was looking at two of these last night, was how specifically the numerator and denominator defined for these components of the outcome and they are specified as DRG's or MS-DRG's and the data source is multiple potential data sources. So you could use claims, medical records, et cetera.

And DRG's and MS-DRG's are deliveries essentially, either vaginal or c-section. But there's also an or statement and I wasn't sure I could follow it completely that identified women who were treated with

ventilatory support I think in a positive way defined to include not just intubation but positive pressure or et cetera. And so I think that the NICU stay component of this may be NICU or, and I just need to clarify with you all, NICU or use of respiratory support.

And I don't know if that's true or if it's just admission to a NICU's, maybe you can quickly clarify. I'm sorry, I have NICU on the brain. ICU.

DR. MURI: No, I think it's an or statement because some of the hospitals you can get at the unplanned admission, through billing data. And with the exceptions it has to be a postpartum complication.

But if the hospital does not have, per se an ICU or do they do not code ICU charges, the clinical conditions were added to capture those women that may be in some other defined area. But have a very serious unplanned complication.

DR. DRYE: So I couldn't totally

follow that. I'm going to sum up and then you can speak to the overall thing. Because there's also in 2A 1.8 there's under denominator details, there's an allowed exclusion for planned ICU admissions for things like placenta accreta, given as an example.

And yet it says under 2B.3.1, which are threats to validity that the planned postpartum admissions to the ICU can be excluded but must be identified by chart review.

So I think there's some, to me it wasn't completely able to discern from the written specifications what is allowed and not allowed to try to pull out planned admissions.

I think it's always good in an outcome measure to try to pull out planned procedures, which those, by definition, are not things, or planned outcomes, these are not, in this case an admission to an ICU. Those are not things that we're trying to

measure or prevent.

But it wasn't clear that the mechanism is really well defined here to do that. Or that medically we could anticipate clearly, and I just don't know enough about OB. What was planned or unplanned and differentiate.

So why don't I sum up and then you guys, if you want to speak, because I know we're running out of time.

So I just in summary, these are ICU admissions, I see days, but it sounds like in hospitals without ICU type care would be captured, the rate is pretty low in this population, 0.1 to 0.4 percent.

You're trying to develop a mechanism to pull out unplanned, I mean, pulled planned out and only capture unplanned. But for me the overwhelming.

So there's a concern for me about the uniformity of that outcome, whether or not someone is admitted or not to ICU given the

way the measure is specified. Have you, you know, really saying ICU admission at the IDMC is the same as an ICU admission in a local community hospital.

You know is there really the same group of patients that you're, it's the same outcome that you're capturing there and I think given out discussion about NICUs and the variation and how they're defined there.

I don't know as much about adult ICUs but I would be concerned if that outcome would be capturing very different things across hospitals. And then just overall the measure is not risk adjusted and I assume that maternal factors are fact a risk in ICU admission.

CO-CHAIR SAKALA: Did you want to comment on that?

DR. MURI: Sure, I think in terms of the definition of the patients, if you have coded data. You can look at antepartum conditions and an ICU admission, which is an

appropriate admission to an ICU.

Versus a postpartum complication, so that's part of the distinction is to look at the fifth digit for the ICD-9 codes. To indicate that this is a postpartum complication, within the range of OB complications.

The other issue is if the hospital does not have an adult ICU, or does not code their financial data that way, charge for that activity. That you have to get at the other level of complications using the ICD-9 codes or procedures. The procedures for intubation, et cetera, et cetera.

The other issue is if a hospital does not have the capability to put the woman in an ICU and needs to transfer that woman out immediately. That's where you get into the transfer issue. Both on the neonatal side and on the adult side.

That the woman or the baby is transferred out immediately, and on the adult

side has a postpartum complication. So the assumption it's not perfect. But the assumption is that they're going out because that hospital cannot accommodate that woman.

DR. DRYE: I apologize, I forgot, I didn't define the numerator. I forgot to say admission to the ICU or transfer to another hospital for ICU admission.

DR. PRATT: I'll just comment quickly on the issue of trying to deal with a plan. You've already heard that in theory there should be a postpartum event, and Janet can get that from the fifth digit.

We tried, we looked very long and hard at whether there were any conditions that might have a high enough planned ICU rate that we would say, let's exclude those. And the fact is there are very few of those.

And we couldn't come up with, you know, mom comes in with an uncorrected tet and happens to be pregnant, it's probably a good idea to put her into an intensive care unit

around the time of her delivery. Extraordinarily rare cause of an intensive care unit admission.

So there weren't any that were a, it's true a procraeto, many of them are going to go. But not all of them and not even maybe many of them. And so the ability to sort of identify that through coding was pretty tough or to exclude specific pre existing conditions. Actually we did not think we could do.

CO-CHAIR SAKALA: Thank you, other members of the committee?

CO-CHAIR RILEY: So I think that this is one of those areas as where you alluded to comparing ICU definitions would be difficult.

Because, I mean, I think, you know, I'm at Mass General and I can watch things on the labor floor. Because I have awesome anesthesiologists looking over my shoulder for a lot longer than necessarily

somebody might do a another hospital.

And I sort of feel like, you know, another bad care would be keeping on your unit to long and not going to the ICU, and not going to the ICU and sort of not recognizing that you got someone who's in shock.

Which I think we hear all the time, so I don't know how to sort of grapple with that. I sort of recognize that the benefit of looking at ICU admissions.

But I just wonder about, sort of, another area where the risk stratification or somehow separating these hospitals, you know, I'm sure that, you know, on Jennifer's unit she can hold on to things a long time too.

And even tertiary care centers are vastly different.

DR. PRATT: We certainly can do the stratification, and that's already been done. And we could do this by, you know, number of deliveries. We don't have the data on things like 24/7 OB trained

anesthesiologists, right?

So that would be more difficult although in theory one could stratify for that if one had those data. The issue of, and this gets, to the previous question of whether people are actually going to keep people out of the ICU in order to avoid this numerator.

We were having such a hard time believing it. Although you also commented on the idea that not, failure to recognize that this patient, where it's not an on purpose decision. You're not saying, oh I don't want to take that hit. But they just aren't realizing it.

Those folks are going to end up in the ICU, or dead, or one of these other events anyway, right? I mean, if they're that sick they're going to get there one way or another. So I'm no worried very much about missing those folks.

CO-CHAIR SAKALA: Other urgent comments? Yes, Kim.

DR. GREGORY: Well, I just, first of all, I guess it goes back to the tenor of the purpose, and that is that it represents bad care.

And I agree that it's important to do case finding, for patients going to the ICU. And the same way that we're looking at maternal events for, maternal deaths as sentinel events.

And the possibility of review systems of care that can be improved, but I would argue that, you know, I would argue especially after reviewing some of the maternal deaths in California that the majority of patients who ended up in the ICU, it was actually a good thing that they were there.

And I'm just sort of I'm concerned that we're sending the wrong message.

DR. PRATT: There's no question that in general going to the intensive care unit is the right thing to do. The fact is

most of the diagnosis that lead to this are associated with bad care. And therefore potentially that ICU could have been.

DR. GREGORY: I would say, for example, preeclampsia is probably is probably one of the leading causes. And that's not preventable. We have yet to be able to prevent it.

DR. PRATT: Not preeclampsia, but the ICU admission might be.

DR. GREGORY: The health, most of them are going because of health, or some TTP variant, or some variation thereof.

DR. PRATT: Those ones you're not going to prevent. The hypertensive crisis you might. They bleed because of the hypertensive crisis. You might. There are lots of, and there's very compelling data that's demonstrating that blood pressure control is done poorly in the severe preeclamptics a very large percent of the time.

Now when I say there's often

substandard care, let me quantify, I'm saying on the order of 50 percent of these cases. So you're absolutely right, and we've talked about this. And this is going to be your problem without come measures forever.

DR. GREGORY: Right.

DR. PRATT: That all of them that I've ever looked at they've got about a 50 percent preventability rate. And many, many, many of those are not able to be risk adjusted at this point. Or identified which ones are in fact preventable.

And you all are going to have to deal with that over and over again, whether that's good enough to help you want to measure it.

CO-CHAIR SAKALA: Last comment please, then we need to take a vote.

DR. DRYE: Thanks, I actually, what's great about outcomes measures from my stand point, is you don't have to know which ones are preventable. You don't, you have to

risk adjust out the patients factors. You have to account for chance variations.

And then you can compare how hospitals do on them and some are going to do better than others. And if there's a lot of variation then we know some are preventable.

And in fact when we think about outcomes measures we try, in my group we don't like, you know, to try to find the preventable thing. You do want to find something you can affect, but the concept isn't can you prevent this and get it down to zero.

It's can you lower the risk environment the patient is in and lower the rate, so that the risk performers start to look more like the best performers.

So I don't think it's an issue whether something's preventable, you know, identifying for any individual patient whether something's preventable or not.

Planned is a different concept, and, you know, it would be wonderful if we

could always tell what was planned and what was not. But the data doesn't support it always.

CO-CHAIR SAKALA: Okay, let's take a vote.

DR. BAILIT: This is a process, not an outcome. ICU admission is a process, it's not a patient outcome.

DR. DRYE: Yes, I think it's on the border, I don't know it's some intermediate. It's not a process in the sense that's it's clinically, you know, it's the standard of care like giving an aspirin at arrival for MI.

It's not that kind of process measure but it's, I'd really call it like an intermediate outcome. I think is what we would call it. It's not necessarily something that's bad for the patient but it's maybe on the pathway to something bad?

CO-CHAIR SAKALA: Okay, can we take a vote please? On unplanned ICU

admission of mothers. Okay, so we have eight yes, and 16 no.

And we'll move on to Maryi for third and fourth degree lacerations.

MS. LESLIE: Okay. So third and fourth degree lacerations, this is not a previously endorsed measure. The numerator is the number of women who suffer third or fourth degree laceration of the perineum during vaginal delivery.

The denominator is all women who deliver. Not all women who deliver vaginally, but all women who deliver.

And with regard to having a high impact it is a national priority partnership goal so it meets that criteria, via that mechanism.

It happens in as high as 30 percent of operational vaginal deliveries and in 2009 that was 3.7 percent of the births in this country.

In terms of there being a gap,

there's several OB management modes such episiotomy and upper vaginal delivery that can be modified that would effect the number of third and fourth degree tears.

And there's some patient risk factors that theoretically could be modified. So it is something that we could have an impact on.

A lot of the evidence presented in the form was really about the AOI, and not so much about third and fourth degree tear but I would say that there's a moderate amount of evidence.

And in terms of reliability, testing at Beth Israel, they commented that there was a strong match in terms of their reliability testing. But they also commented that when there was not a strong match, and they described what they did about it.

So it suggested that maybe there was some problems in terms of the match during the reliability testing.

And I think one of the things with third and fourth degree tears is that it's been documented in several studies that there's a lot of variation between provider diagnosis and also coding. There's problems with coding.

And it's mentioned even in the form that like a second degree tear with an extension is coded as a third degree tear. And a third degree tear with an extension is coded as a fourth degree tear.

And there's some variability in how it's coded. So reliability might be not as good as it could be.

Validity testing based on what's said in the form I think is high. And it has been used for public reporting both by JCo and by AHRQ and I think it appears to be easily collectable data. I'll open it up.

DR. BAILIT: I wish Bill Grobman were here, but he's got a beautiful study from 2006 in the Gray Journal that talks about why

this measure has to be risk adjusted and is not a great quality measure. So I refer the Committee to that paper.

CO-CHAIR SAKALA: Comments from anyone? Kathleen.

DR. SIMPSON: I just wanted to mention that the joint commission decided not, that that measure was retired. So it's not something that is ongoing, that was retired from the first. And then the incidence of episiotomy was decided would be a better thing than a third and fourth degree. That's not current.

DR. WATSON: What's substandard about getting a third degree midline laceration on a vaginal delivery of a nine pound baby? What's substandard about that?

DR. PRATT: I'll let the obstetricians speak to that so, Susan?

DR. MANN: One of the adverse outcome is that, I apologize for the echo, was created it was at a time when both joint

commission and AHRQ were collecting this data.

And it was felt that obviously it was important to bodies higher up than the consensus panel with the results that it needed to be included.

And that was the thinking then because the literature has been, or measurement has been done historically with it that has remained in the adverse argument there.

DR. WATSON: But that's now been retired and we're just looking at episiotomy rates. So the emphasis has changed.

DR. PRATT: To answer your question, probably nothing. Although again, I'm an anesthesiologist. The second question though is why are we looking at episiotomy rates? Or why are you all looking at episiotomy rates? And my assumption is because it's associated with an increased risk for third and fourth degree tears.

You're right, so what we're

getting at is, if you do an episiotomy and nothing bad happens and the provider gets dinged, and it speaks to now, are we looking at measures, that are process measures and as provider, I love that.

Because I can give my antibiotics within 60 minutes of incision, and I've done a good job and our hospital looks great on our SCIP measures, and off we go.

But if the patient gets a wound infection despite that the patient does care how well I did. The patient got a wound infection. And if we look at care from the patients stand point, rather from the providers stand point.

They don't care whether or not you cut an episiotomy they care whether or not they got a fourth degree tear and now have a fecal incontinence for the rest of their life. right? That's what they care about.

And so that's why the outcome. Now it is true that we've kept this largely

for historic reasons and we've thought about it over and over again. And it's entirely possible that we'll end up retiring this as well.

But still, philosophically, you all are saying let's change this so that we measure what the clinicians are doing and I'm saying I care more about the moms perineum than the decisions made by the clinicians.

DR. WATSON: But I think the point is --

DR. MANN: The performance of episiotomy would not be a third and fourth degree, you know, when it's coded, it's coded as an extension to the third or fourth degree.

CO-CHAIR SAKALA: Nancy, did you have a comment?

DR. LOWE: Yes, I guess we have to remember though that for a huge two, three generations of women. They regularly had their perineums incised. And I can tell you it was not pleasant. And I can also tell you

that when you had an extension of one of those it was not pleasant.

And what we do know from the literature and from the data is that episiotomy is the precursor to the third or the fourth degree extension. That's so clear. That that's the precursor event.

And so that's why we've already adopted that as a measure. And so to look at simply third and fourth degrees does not, to me, does not get at the sentinel event. If we want to use that terminology. Which is the episiotomy itself.

And once you've done that, like someone said, I don't know if it was Laura that said this, someone said. You know, if you're, if it's a difficult delivery because it's a large fetus. And we've done the episiotomy we've already set up the situation in which the woman is likely to have a third or a fourth degree lac.

So, you know, I don't buy your

logic that the episiotomy is not the appropriate thing to measure. Which is very easy to measure off the chart.

DR. PRATT: From a measurement stand point, you're absolutely right, it's easier. And we're in agreement that that the way that the third and the fourth degree tears would be influenced is by creating practice improvement around episiotomy.

Also probably around operative vaginal deliveries. The OB anesthesia literature demonstrates that there are many obstetricians who do operable vaginal deliveries because moms got an epidural in. Not because she necessarily needs it but because she happens to have one in.

Or this is the one they teach their residents on because she's got one in. So there are lots of places in both of those where there could be quality improvement efforts to decrease the rate of this outcome.

As I said, it's a question of

whether you want to get at measuring the risks and the process level or the actual outcome. And that's somewhat philosophical.

CO-CHAIR SAKALA: So, Kathleen, and then we'll have a vote.

DR. SIMPSON: I just wanted to mention that if you did go from episiotomy to, I mean from third and fourth degree to episiotomy, right now third and fourth degree laceration already drives this measure.

When it's the most common thing, is my understanding. So if you went to episiotomy then the whole thing, AOI, would be about episiotomies because then you would have a higher number of those.

Rather than, you know, it would totally screw up the whole thing. I mean, that's all it would be about is episiotomy, I agree with you that episiotomy is the bad event.

But I'm just, you know, from the purpose of what they're trying to do is that

would be pretty much, it would all be about episiotomy.

DR. LOWE: I was objecting to the idea that women don't care. If they have an episiotomy. That's what was said. Trust me, we care.

(Simultaneous speaking)

DR. PRATT: Given my gender I will do better about trying to presume what women do and do not care about. Fair enough.

CO-CHAIR SAKALA: Thank you. Let's take a vote. So we have seven yes and 17 no.

Now we've just made an executive decision that the way we're going to pick up some time here is to have a working lunch. So I'm going to ask you, it's sandwiches, easy to do, to five of to be back, okay? So don't get sidetracked with conversations, et cetera, phone. Thank you.

(Whereupon, the above-entitled matter went off the record at 12:52 p.m. and resumed at 1:05 p.m.)

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

1:05 P.M.

CO-CHAIR SAKALA: Okay. So we're going to move ahead with Lee, on Unintended Operative Procedure.

MS. PARTRIDGE: This component addresses so-called Unintended Operative Procedures. The numerator is the number of women who, during their delivery/hospitalization, which for me turned out to be important, have an unanticipated operative procedure. Basically D&Cs or repair or control of hemorrhage.

There are a few more codes in there but that's sort of the broad categories. The question that I had as I read this through was, and again, an awful lot of the material that was submitted related to the whole AOI and not necessarily specifically to this component.

The fact that you're tracking this only during the actual admission for the

delivery suggests to me that you might miss a significant number of cases in which there was, somehow the case was not handled as well as it might have been. And the woman has gone home.

And then the hemorrhage develops or the need for the D&C or et cetera. So I was just curious as to whether or not you all felt comfortable that this was adequate as a measure of what were perhaps inappropriate care.

DR. PRATT: Dr. Mann.

CO-CHAIR SAKALA: If you're on mute we can't hear you.

DR. PRATT: I can speak to that a little bit. The issue, of course, there is the capture of those who have gone home. We can of course capture them if they came back to the admitting hospital.

But if they went somewhere else we would be unable completely to identify those folks. And that's why this decision was made,

primarily.

MS. PARTRIDGE: But as I read the description, even if they returned, if it wasn't within the original hospitalization it wouldn't be in your numerator. Am I wrong?

DR. PRATT: Correct. The reason to exclude those is you end up now speaking to different practice areas. You know, I'm in Boston you throw a stone, you hit four hospitals. Right?

Whereas you're someplace that has only one hospital for four miles they are going to get all of those returns. And I'm likely to get none.

And so it wasn't reasonable to have those come back and count in the numerator when the practice patterns would be so different from one location to another.

MS. PARTRIDGE: In essence the damage could have been created at Hospital A and the person was admitted to Hospital B for the correction?

DR. PRATT: Right.

MS. PARTRIDGE: Okay.

CO-CHAIR SAKALA: Other comments?

DR. DRYE: So a downside of that strategy, because I think that's a good example where you thought about differences in systems in defining the outcome.

But the downside is that hospitals with shorter lengths of stay, who send their patients home earlier, which may put them at risk for the outcome they're going to get let off the hook for that or be held even less accountable.

MS. PARTRIDGE: Well I would probably also add that sometimes it's not the hospital that determines whether you stay a short time, it's the patient herself in this instance.

CO-CHAIR SAKALA: Other comments on unanticipated operative procedure? Okay. So let's take a vote then. Okay. So we have 15 yes and 9 no.

DR. PRATT: You're just happier now that you have food.

(Laughter)

CO-CHAIR SAKALA: Okay. So Sarah Brown's not here.

DR. WINKLER: The next one is Measure 750, which is Maternal Blood Transfusion. And we do have the DRG and the MSDRG codes for blood transfusion, again, as a part of the component.

I think there's been some conversation earlier about appropriateness of blood transfusion at times. But I think blood transfusion in and of itself is fairly straightforward.

CO-CHAIR SAKALA: Any comments about Maternal Blood Transfusion as a component of this composite measure? Okay. So let's vote. So we have 17 yes and 7 no on that one.

And we're turning to Five Minute Apgar Less Than 7. Angelio.

DR. JALEEL: So I had difficulty navigating through the submitted document. Partly it's because it does not address the specific component.

CO-CHAIR SAKALA: Could you speak a little closer to the mic? Thanks.

DR. JALEEL: I have my plate up here. So I had difficulty navigating through the submitted document. Partly it's because it does not specifically address the specific component. It addresses the general measure.

And partly it is because the questions are not specifically answered in multiple places and things like that.

So anyway, the measure is Five Minute Apgar of Less Than 7. And my thought is that, is a low Apgar a measure of substandard care or a preventable complication? So that's the issue here. So one of the references that has been mentioned is the citation for evidence of high impact is the policy statement by the American Academy

of Pediatrics.

Actually when you read the AAP statement it actually focuses on being more cautious about the use of Apgar score, because a number of factors may influence the Apgar score.

It is not appropriate to use Apgar score alone to establish the diagnosis of asphyxia, because there are so many other components.

Maybe a better measure would be checking of cord blood gas for fetal acidemia so that would probably be more appropriate.

And an Apgar score at five minutes, in term infants, correlates poorly if you are looking at the neonatal outcomes, it correlates poorly with the neurological outcome.

So that doesn't help either. And can it be used for monitoring the delivery service? Yes it does say that it can be useful.

But with the caveat it says specifically in that document that individual case reviews are more, can identify these issues and it probably doesn't help as a quality measure for, what's the word, I forget. For disclosure.

Yes, public reporting that was the word I was looking for. So I'm not sure whether this is a good measure or not because there are multiple factors which do factor in.

And one of the citations for data on the performance gap that they have mentioned are three citations, which specifically again, do not address Apgar score but address the index or such.

So it doesn't help in that. And then there's one more citation which looks at a Swedish registry and looking at low Apgar score at five minutes.

And I went back and looked at that article and it's a Swedish registry. And what weighs in more is the birth weight and the

gestation.

And there are several observed risk factors which are associated with the low five minute Apgar score in term infants. But many of them are preeclampsia and things like that, which we don't have control on. So I'm not sure whether it helps either.

Maybe as a composite measure it might help. But individually is this outcome measure an important one which can be reported publicly? I don't think so.

DR. PRATT: Just quickly, we do restrict this to term babies at 2,500 grams, so we're dealing with much of the prematurity issue. And yes, as is true with all of these, there are many factors that can influence it. But certainly does include obstetric management and obstetric decisions.

And we also exclude many congenital anomalies to try to look at basically healthy term babies who now come out with a pretty low Apgar score that is

maintained.

It's not just the baby that has a vacuum or assisted delivery and comes out and is a little stunned and by five minutes it's pink and screaming.

These are babies that carry on to have a fairly low score. So we tried to narrow this down to babies that really should be good, you know, good babies by five minutes.

This isn't trying to measure whether they're going to have high grade HIE, it's not going to assess whether or not they're going to get into Harvard when they're 18.

But we tried to, as best we could, find just those babies that really should be doing well.

CO-CHAIR SAKALA: Other comments? Kathleen?

DR. SIMPSON: I just had a question on the references. You said that

about 50 percent of the low Apgars at five minutes are related to intrapartum events. But there's not reference cited for that.

And so I, like Jaleel, was looking for citations specifically related to the low Apgar. And that one really was interesting because that does give you some idea as to maybe it could be prevented or not.

But I didn't see the reference. So could you provide, or do you know where that's from?

DR. PRATT: Dr. Mann wrote this up, wrote this particular one. We all sign off on all of them. But she was the one getting those references. Susan, are you there still?

CO-CHAIR SAKALA: Okay.

DR. JALEEL: Yes, that's a concern with many of the statements which have made. There are no references or the references are in different places, which don't actually, it's not well documented.

MS. PARTRIDGE: I found that a challenge, I must say, in reviewing all of this. Some of the references are incorrectly cited too. They have like the wrong volume or the wrong year. I mean I eventually found most of them, but that was frustrating.

And then the just broad range comments that this is related to substandard care, no reference. I think it could have been better prepared I think.

DR. PROFIT: I'll tell you when I was putting references I cut and paste them from PubMed, so if there was an issue with which volume we can talk to them. So the ones that --

MS. PARTRIDGE: Okay.

CO-CHAIR SAKALA: Any other urgent comments on 5 Minute Apgar less than 7?

DR. PROFIT: I guess I was just wondering why you didn't use Apgar at one minute rather than Apgar at five, because most of the index seems to relate to kind of like

prenatal, like obstetric care.

And generally lest we kind of conceive of the personal doctor's score consistent with maybe prenatal care versus five minute Apgar scores also then dependent on resuscitation and technique of resuscitation.

DR. PRATT: Certainly that is true. But there's so many babies who, for various reasons, are born and have a relatively low one-minute Apgar. You know, precipitous delivery can cause a low one minute Apgar.

And we certainly wouldn't want to put that into the numerator. And the other part of that is, yes, again remember this is looking at the overall quality of care, not just the issues related to obstetric management.

And so, you know, who's in the room. Are there obstetric nurses doing all that resuscitation? That speaks to the

overall labor and delivery care.

Just quickly speaking to using cord gases. It may well be a better measure. It certainly is a better measure of acute acid base. One of the issues there is remember the obstetrician is in charge of whether or not that gets sent.

And there certainly is the potential for some bias there in terms of whether or not they send it. My guess is that most are sending it to be defensive and over-send. But you could certainly imagine them not wanting to know what that pH is.

Versus most of these Apgars are being measured by people other than the delivering obstetrician. Pediatricians, nurses, et cetera. And so it is a measure that is by somebody who is not as invested in that delivery.

DR. PROFIT: Apgar score, as we know, it is a very subjective measure. And there have been multiple studies looking at

multiple providers, neonatal providers and pediatric providers doing this and having variable results. So it's not a very good measure.

And you can probably mandate and say if you have these risk factors then you should get a cord gas. Is that something that you can do? I'm not sure, but I --

CO-CHAIR SAKALA: Okay.

CO-CHAIR RILEY: That is actually in guidelines for perinatal care. With Apgars less than X you need to get cord gases. And you know, we're at CRICO institutions, if you don't get it they're not helping you out when you have to sit in the big room. So I don't see that, actually, as an impediment.

I think the time that people are should I send them, should I not send them, the Apgar is good but the tracing was bad or something else was going on.

And you're thinking do I really want to hang myself them. I don't thing

people game that system in the wrong direction actually.

I think if anything people are getting more information. And it may be medical/legally motivated. But nonetheless I think that they get the information.

CO-CHAIR SAKALA: Okay. Could we please take a vote now? Suitability of 5 Minute Apgar Less than 7 as a component of this measure.

So we have 10 yes and 14 no. And we'll turn to Nancy for Birth Trauma.

DR. LOWE: the Birth Trauma Measure, if you remember from our discussion yesterday, it is very similar to 474 that we talked about yesterday, that we did not endorse.

There is one primary difference between 474 and this measure, in that in this measure brachial plexus injury is included. So I'm not sure we need to revisit all that stuff, because it's the same measure.

The weights are the same. The ICD-9 codes are the same, except for brachial plexus injury, which is included in this measure.

MS. WATT: I just wanted to mention, there is one other code that is different from the PSI 17. 767.8 is not included in this measure, in this metric. 767.8 is Other Specified Birth Trauma.

The other difference is that the denominator here is deliveries, obviously because it's part of the AOI.

CO-CHAIR SAKALA: Are you done, Nancy?

DR. LOWE: I mean, I just don't think there's a need based upon the long discussion we had yesterday about this.

CO-CHAIR SAKALA: Okay. Anyone else? Okay. Let's take a vote on this. Birth Trauma as a Component of the Index.

Yes, others who haven't weighed in. Okay. So we have 8 yes and 15 no. So

we'll turn to Bill Callaghan for In-Hospital Neonatal Death.

DR. CALLAGHAN: Thanks. I have to admit I was a little confused by this measure as specified. In the numerator it specified as any in-born, discharged disposition of died within seven days of birth (perinatal death), which really isn't a perinatal death.

A perinatal death is an early neonatal death. So perinatal death would include still fetal deaths. And then in the AOI papers they specify intrapartum and neonatal death.

And more specifically these are greater than 2,500 grams and more than 37 weeks term. So I took this on face value to be early neonatal deaths. That is born alive and died within seven days of birth.

Okay. Obviously, again, this is another rare and sentinel event. There are data that were published in the Journal of Pediatrics in 2007 that actually looked at

this, in a different context, to compare late pre-term births with termed births.

And if, as specified in this definition, you eliminate congenital malformations as a cause of death that would limit this to about 750 deaths a year.

About one in five of those deaths would have been assigned a cause of death as birth asphyxia or intrauterine hypoxia and then the rest various and sundry causes of death.

So this represents a fairly heterogenous and rare group of deaths. It's difficult to understand for me, where sort of the, again from an accountability standpoint, who is accountable for what.

If the intent was that this was about intrapartum management then I think it gets even sketchier, because again this is about 750 deaths per year in this group of non-anomalous term fetuses. One in five of which would have been assigned intrauterine

hypoxia and birth asphyxia.

And it still remains unclear, I think, in this day and age how much of that is really and truly intrapartum management, how much of it isn't.

From a citation standpoint, there's some citations from some international articles that have more of an international focus where intrapartum death certainly is a very different phenomenon that it is in the U.S., as well as asphyxia.

Some references to infant mortality and neonatal mortality and post-neonatal mortality in general. But not specifically addressing the issue of early term neonatal mortality.

So I think, again, at the end of the day, from my perspective, this represents a group that is pretty hard to hold out again for accountability or preventability.

CO-CHAIR SAKALA: Comments, questions, from anyone around the table?

DR. PROFIT: I guess for me it's interesting it took a group of obstetricians to give the neonatal death only about half the weight of maternal death, but a point well taken.

DR. CALLAGHAN: Was the intent basically based around intrapartum management? Because the focus of this, although it's been unstated I think explicitly, is that this relates to intrapartum care.

DR. PRATT: Yes, absolutely. This was designed to be around intrapartum management. There have been several other studies looking at entire pregnancy management and various outcome measures associated with managing the entire pregnancy.

But the entire AOI has been about measuring quality of care in labor.

DR. CALLAGHAN: But not, for this measure specifically, not care after delivery?

DR. PRATT: Not, so correct. So one of your points may be --

DR. CALLAGHAN: And if it dies before seven days of life may have been of the Apgars.

DR. PRATT: Yes, might be his fault.

DR. CALLAGHAN: Probably was.

DR. PRATT: Fair enough. And you're right. We did not include this as a measure of nursery or NICU care.

DR. PROFIT: I'd like one more comment, I think goes back to the stratification by kind of hospital. And maybe I missed this in here but in a lot of smaller hospitals a baby may be admitted to a NICU or Level 2 nursery or something that you would define as a NICU but then get shipped out to a tertiary care center.

And so I couldn't really see how, or didn't quite understand how this would be captured in your index. So they may not die in your hospital. There's some hospitals that pretty much don't want kids to die there and

they'll end up dying somewhere else.

DR. MURI: Yes, I think that's a real problem with any measure at any hospital that has the opportunity to transfer out the problem. You know, you're not going to pick that up unless the receiving hospital gets dinged for it you're just not going to get it.

But you're not going to get it from the hospital that's doing the transferring, unfortunately. Better linked data.

DR. PRATT: Yes, although we do, for the neonatal ICU admission, we do get those transfers. So we would still identify it as a case. We likely would miss the death. And I think that that's going to be true for, it was also true for the unplanned procedure, right.

But at least we would still get this as an event because it would have been transferred out, likely to go a NICU and we do capture those.

CO-CHAIR SAKALA: Okay. Can we have a vote then on In-Hospital Neonatal Death as a component of this measure? Okay. So we have 8 yes and 16 no. And Elizabeth, we turn back to you for the final component. Admission to NICU at Term.

DR. DRYE: Okay. So this is our last component of our last measure, I think, until we go to the harmonization discussion.

The measure is for babies who are at least 2,500 grams and 37 weeks who are admitted to the NICU within one day of birth and who stay greater than or equal to one day in the NICU.

Or who are transferred within one day, which is parallel to the maternal ICU admits. The challenge with this measure is, well the baseline rate reported in the application is six to eight percent. So it's much higher than a lot of the other things we're looking at.

And the challenge is that, again,

the two big challenges are are that there's not risk adjustment and I'm sure that both maternal and infant factors contribute to their NICU admission risk.

And then thinking about that as a threshold, admission to a NICU or transfer, there is, as we've talked about already so we don't need to go over it again, a lot of variability of the rates of admissions to NICUs. Some of which may be a quality signal.

Some of which may be a structural or inefficiency signal that, you know, if hospitals are admitting 25 percent of their babies to NICU that may be a different kind of signal that we would be getting out of this count that would be feeding into the composite score.

The information on this specific form for this sub-component is pretty brief and I don't blame the metric developers, it's a lot to fill out ten different forms for all these outcomes is challenging.

But there's very little on the validity of the measure. The studies cited are really back to the index, not this individual component.

There is some look at the data validity and there's subjective words like, you know, there was a strong match or whatever.

But there's not quantification in the application about how well the different data sources lined up in determining who was admitted to the NICU or not. So I think those are the main points.

CO-CHAIR SAKALA: Other comments, questions?

DR. MURI: I just wanted to mentioned that one of the ways that we try and do discriminate in terms of the processes at a hospital is that the baby actually has to be in the highest level bed.

So many hospitals will have intermediate level beds. They'll have

step-down beds, they'll have observation beds. So these babies in this 2,500 gram, 37 week can be taken care of for a couple of days in those types of beds. But if you're putting a baby, without the exclusions, if you're putting a baby in a NICU level bed then there are other drivers potentially that are causing that. So, I think, that's why we're trying to discriminate that the baby can't be just in a special care bed, but it has to be in a NICU bed. And so that's why you get those very large numbers of cases.

So when hospitals begin to look at their data they start saying, oh, they really shouldn't be at that level. They could be in an observation bed for rule out sepsis and those types of things, so is it really appropriate care.

DR. DRYE: So let me just clarify one thing, which I forget to mention. This two out of two numerator statements I've incompletely presented to you. So I

apologize. But this one does have an exclusion for babies with congenital anomalies, fetal hydrops, dwarfism or neonatal abstinence syndrome. And codes are provided for those.

I'm not sure what you just discussed about level of bed addresses the observation we were making earlier that some NICUs, say in a community hospital, there may be a Level 2 NICU. And it's really the only place where there are nurses and potentially a on-call resident, for example, this is my experience in a small community New Haven hospital, really able to care for babies where there's any concern. And so the threshold for admission is very low versus say at a tertiary care center like Yale. So I think that's a difference that we're all acknowledging.

And the issue is in a composite measure like this, what would the signal be giving you. You'd be capturing potentially two very different things. One is sort of an

inefficiency in the system or a management choice that we don't think makes sense versus a quality signal that was poor perinatal care during the delivery.

DR. MURI: I think for that type of hospital if the baby really needed the NICU level bed that baby would not be picked up on this measure. But if the baby really needed the NICU level they would be transferred out and they would be covered under that transfer scenario.

DR. DRYE: So let me just clarify then, because I'm confused. If the hospital, the highest level is a Level 2 NICU bed, say a six-bed NICU, would you count that? That would be a NICU admission I think documented in the charges.

DR. MURI: A Level 2 bed NICU is driven by the revenue codes, the UB-04 revenue codes. So a Level 2 bed is coded at 172 or 173, the only ones that we're picking up are 174s. So it has to be truly charged at the

NICU level. They may call it a NICU.

DR. PRATT: At the Level 3 NICU level.

DR. DRYE: Okay. Well that's a clarification. I think it's not really specified in the form. And again, these forms are kind of endless and they're hard to do.

And let me just say reviewing it, it would have, I think, it would be helpful just to have those kinds details there and also even just the DRGs listed so that you don't have to look them up as -- just a description of the DRG not just the number so you don't have to look them up to understand the basic things like numerator and denominator statements.

CO-CHAIR SAKALA: Are we ready to vote? Thank you. Admission to NICU at Term as a component of the composite measure. Okay. So we have 11 yes and 13 no on that.

So Suzanne has prepared a summary slide of our votes. She's just putting in the

last one and then we'll put it up. We had a tie, a no, a no, a no. A yes, a yes. A no, a no, a no and a no in terms of the simple majority votes. But now we're going to ask you with no further discussion? Yes. Okay.

So we're going to go through all of the four major NQF criteria, thinking of this composite measure as a whole. So the first question is importance to measure and report. Opening it up for comments.

DR. GEE: I just wanted to ask, and Scott and I were just talking and we both had the same question. On a national level certainly the maternity field is very premature in terms of our measurement science.

A measure like ths is extremely complex, certainly now with, we don't know I think now, four of the ten are still in there. How do we vote on something like that where we don't know -- understand -- okay, if you redid it with only a few measures and you want to weigh it a certain way, how could you think

about it?

Then on a national level who else are using composite measures? Are they well validated elsewhere? Are there other professions that use these, and what are examples of those?

DR. WINKLER: Yes. Two questions. First one is you need to vote on what's presented. So there's no tweaking, you've given a lot of feedback, but that's not what you're voting on. You're voting on what's right here in front of you. Okay?

So the second. Are other composite measures used? Yes, they are. There are two basic measure kinds of composite measures that we've endorsed. We've endorsed, several now, what are known as all or none composites where there will be multiple components.

Often process measures, but they can be outcome measures. Where it's, for instance, good diabetes care means your

hemoglobin A1c is less than whatever your number is. Blood pressure is less than. And, you know, the LDL is less than.

And the patient, in order to get credit for the measure, you have to hit all four. Okay? So that's a kind of measure we're seeing a lot more of. And we've endorsed any number of those and they are being used nationally.

The other kind are the kind that Elizabeth's worked on. And these are more traditional composites where there's weighting and risk adjustment of the components and factor analysis and, you know, they all get kind of put together a little bit odd.

(Off microphone comments.)

DR. WINKLER: Microphone, please.

DR. DRYE: Both are outcomes measures that we've -- one's just finishing up the NQF process and one's approved. They are risk adjusted but they are specified in the way that these are. It's one or more of six

or seven outcomes. And so we didn't treat them as a component per se. We looked at each of those clinically and at the coding variation and asked a lot of the questions that we talked about today, but they weren't true composites.

And I think, you know, everything is evolving.

DR. WINKLER: Yes, that's one of the evolving issues, is when is a composite composite and when is it something that there are just multiple criteria to meet for the numerator. And at this point it's sort of arbitrary and we're still trying to find our way through that.

But there are those kinds of measures endorsed and being used.

CO-CHAIR RILEY: Getting back to that, that two. One of the main questions that's been brought up all morning is how would the public use this. Like it's complicated. So how would the public use

this? So I guess my question is --

CO-CHAIR SAKALA: Are you at usability or -- ?

CO-CHAIR RILEY: Oh, okay. All right I won't ask the question that I was going to, shoot it back to Elizabeth.

(Off microphone comments.)

CO-CHAIR RILEY: So she told me to stop talking because --

CO-CHAIR SAKALA: Other comments on importance to measure and report?

CO-CHAIR RILEY: For the composite?

CO-CHAIR SAKALA: Yes, for the overall composite of all ten components. Okay. Please vote. Okay, so the vote is 7 yes and 17 no. And now we do stop according to the process, because the first two criteria are must-pass criteria. But we thank you for your work and your time.

CO-CHAIR RILEY: So now I do get to talk? So one of the major issues,

obviously, that we're having with this is if you were to put this out to the public and they're going to compare hospital to hospital, you know, we're getting really, you know, having chest pain thinking about it.

So with those composite measures, or whatever you call the measures that you're doing, how is the complexity of those rolled out to the public. Like, what does that look like? I'm just trying to understand.

DR. DRYE: On the two that we've done they're not publicly reported yet. But I will just tell you I think the way we'd handle it would be the same as the other outcomes measures. So there are one or more of five or six complications following ICD implantation and the other one is similarly.

I can't remember off the top, I think it's six or seven different complications following elective hip/knee replacement. So what would show up is that there are one or more complications.

The drill down data would got to providers so they could have that. And it's supported with a risk adjusted rate, but also with an uncertainty estimate around that rate, which is challenging enough to communicate to consumers.

Let me just make one point I thought about earlier with respect to the usability of this measure. You know, it's just worth thinking about when do patients have a lot of choice.

So for example, we developed CMS reports AMI mortality and AMI readmission risk adjusted rates on the website Hospital Compare. And we don't think of people really choosing their hospitals by looking up, you know, that data online when they're having chest pain, speaking of chest pain.

But it's, I would think a little bit more and you all can tell me more. I know there are payment and provider arrangements with specific hospitals. But the patients

might be more sensitive to using this kind of measure to make choice.

And so in that sense I think we have to be judicious about putting, you know, we want them to have more information but this is one that's more likely to get used for patient choice.

DR. WINKLER: I'll take the prerogative, I'd like to introduce the Senior Vice-President of Performance Measures here at NQF, Dr. Helen Burstin, who snuck in earlier today and she'd like to make a comment.

(Off microphone comments.)

CO-CHAIR RILEY: So I just wonder if, and that makes complete sense to me, but as I think about it people are coming in all with one thing.

So you had a hip replacement, so everybody had a hip replacement, and then you're going to stratify based on, you know, sort of how you are as a patient or your age or whatever it is or how sick you are.

And that to me is straight forward, like we're all talking about the same thing. And a little bit of this is that we're talking about vaginal delivery, we're talking C-Section.

And it's not as homogeneous. Like I don't know whether it would be easier to understand if we were talking about everyone having a vaginal delivery and then, you know, sort of composite measures or whatever versus -- No?

(Off microphone comments.)

CO-CHAIR RILEY: Okay. Their expectations are very different.

(Off microphone comments.)

MS. PARTRIDGE: When I first saw this measure I was very excited. And I don't want you developers to go away feeling, I'm sure you're going to get a little discouraged at this reaction.

But I was excited for exactly what Helen was talking about. This is one of the

situations in which a woman, or a family, as sometime's it's a joint decision, can get to choose which facility she goes to.

And I'm one of those people who does read Hospital Compare. And when my husband was contemplating a hip replacement I read all of the stuff that I could find about the Washington D.C. hospitals and I think it considerably narrowed his choice of both surgeon and facility.

As we've worked our way through this measure today, as you can see, there were a number of concerns about a number of the components.

And I think part of it is the problem that we're used to dealing with measures individually. Judging each particular measure on its own bottom.

Here we're trying to do a market basket and that's a little harder to get your head around. But please come back at some point when there are opportunities to do that.

Composites are something that we do welcome.

And I do think they can be explained very easily to the patient. And I think, frankly, it's the kind of information that you're attending, your docs and nurse midwives too, because in Washington we have both, in Washington like to know.

If you've got a sloppy hospital you'd like to know it. You don't want to be associated with it, or you'd like to see it get better.

DR. ARMSTRONG: You know, I would say again echoing the payer perspective, sort of the first introduction of this is very attractive.

A huge interest on the part of payers, an enormous interest on the part of planned sponsors, companies that are paying the bill.

But it puts a great responsibility on the developers to make sure that these are really the markers of adverse care.

And that the weight attributed to each of these things really reflects quality and the opportunity to change it, because as soon as it goes out into the marketplace the plans, the companies, the plan sponsor's employers will say, ah, this system has a 450 score and this one has a 320.

I want all of my employees to go here. And I want you to create financial incentives to make sure that happens. So obviously that's how it's used.

DR. DRYE: I just want to follow-up on Laura's point because you made the point and we just want to distinguish between how we define the cohort of patients we're measuring and then how we define the outcome or a set of composite outcomes.

And so for hip and knee surgery, for example, those are elective procedures. We clunked them together because they're done by the same surgeons in the same parts of the facility with the same support team.

And then we looked at a set of outcomes that we could combine together as, you know, consistent with your strategy to get enough volume to really be able to risk adjust for patient differences and still get some variation that wasn't just due to chance. That could be attributed to quality.

So the dilemma you might be able to take, you need to have a sum coherence in the cohort you're measuring. If you're throwing in sort of high-risk OB patients, and you know who they are in advance. And they know who they are.

You know you might stratify some OB patients and look at them with a composite outcome and then have sort of younger, healthy patients without any comorbidities in another group and look at a composite outcome for them.

And what you do by doing that is you get more out of your risk adjustment. I mean, technically, it's what you want to do.

You want to have clean cohorts. That's part of the sort of standards for outcomes, measures and risk adjustments.

So it's not an either/or. What we're talking about that consumers like is the composite of sort of events at the end. It doesn't mean that you have to smush all the patients together who are different, you know, in order to do those composites. I'm not sure I'm being clear.

CO-CHAIR RILEY: Yes, that was incredibly helpful, because it's sort of, it's not that you would think, oh yes, I'm going to have a C-Section and I want to be in this box, because obviously you're thinking I'm going in to have a baby.

But I do think patients are pretty sharp, they know when they're high risk. I mean sometimes they know better than we do.

But, I mean, I think that's an easier way for me to think about it in terms of if you're a low risk person you're thinking

man why does that hospital have so many people that are getting blood transfusions. What the heck is going on there?

And I would be very suspicious if you saw that a lot of those low risk women were getting blood transfusions. So maybe that's another thing to consider. That I can understand.

DR. BAILIT: So I think this measure's still got life. Here's what I would suggest, because I think this needs to be risk adjusted. I think there are a couple of ways to do it.

There are a sophisticated risk adjustment models. As we've seen with C-Section they've kind of fallen by the wayside. I would do AOIs only in the same population you're doing the nulliparous term singleton vertex rate.

And by narrowing the population, and if you stick with the same population the NTSV rate's already doing, the hospitals have

already figured out who those people are, every hospital has low-risk women. Don't hold them responsible for the high-risk.

But what's happening in that low-risk population. I think that would be an easy way to resurrect this.

DR. PRATT: First I want to say thank you all, this has been very interesting, very fun for me. I'm an anesthesiologist so I like getting beat up. Surgeons do this routinely.

The questions I was going to ask were, sounds like risk adjustment. And it sounds likes perhaps even you'd be happy with risk adjustment into some rather large categories.

So low-risk/high-risk women, perhaps. And maybe even, I'm not sure about cesarean delivery versus vaginal delivery. Yes. But certainly even if we were to say low-risk/high-risk, even that would make you happier with this. That's one.

Second, potentially, is stratifying by levels of care and getting some of that more clearly delineated. We've done some of it but present some of those data and make that more clear.

I know you guys had some questions about the weighting of each of the, and we could certainly go back and, you know, this was done by consensus and we're all going to, you know, if we asked all of us around the table my guess is we'd get 43 different answers and there are only 40 of us in the room kind of thing.

And then finally, there were a couple of them, while it seemed to me like most of these ten you had issues with you thought there was a reasonable measure potentially, but wanted risk adjustment with it.

There were several that you had real concerns about and I would say that that's third and fourth degree tear or

lacerations. And Five Minute Apgar were the two that it sounds like just basically you had issues with at a basic level of the measure itself.

CO-CHAIR RILEY: I think that certainly for Apgar it was the science, I think, was the sort of really measuring the --

DR. PRATT: Right. That subject and the third and fourth degree tear, I'm still going to internally push back to do we do this. This measure is about to be about to be about patients rather than processes of care. So we'd have to have that conversation ongoing.

MS. PARTRIDGE: I think that the third measure that got an overwhelming number of no's was uterine rupture.

(Off microphone comments.)

DR. PRATT: Oh yes, right if we did lower at risk versus birth, I mean, there wouldn't be very many. It would likely go away. The vast majority of those either get,

if we care about them, they get a transfusion or go to the unit anyway so we're likely to find them in other ways.

CO-CHAIR RILEY: And if you've got a uterine rupture in that situation you ought to look at that. It really is rare, rare, rare, rare. It does happen, but it's rare.

(Off microphone comments.)

DR. SIMPSON: I just wanted to say one thing. I like the way that you mentioned that it was sort of a one subject. And you said hey, this is really all about labor care. And maybe you could repackage this and the concept and say that it's about care during labor. I never heard that until you said it.

I read all the stuff about it and I seen this thing and that thing and all over the map. And yet, when you said we're really looking at care during labor. How safe is it during labor.

Maybe if you somehow focus that and that was your message, and maybe even

change the name of it, you know, you might be in better shape.

Well, because if that's what you're trying to measure, is safe care during labor, that makes a lot more sense than all the rest of this.

DR. PRATT: I have one more question, I know you guys have lots of work to do. There was one sort of opinion piece on this that suggested that we turn it on its head and do one minus the AOI in essence. Suggesting that what consumers might want is what percentage of moms go in and have a good healthy delivery.

And therefore you would say, here we go, 93 percent of these moms came in and had nothing bad happen to them. As opposed to here's the percent that had something bad happen to them, if you will.

DR. GREGORY: I can refer you to a paper, where I'm the first author, called the Ideal Birth Outcome. And so that has been

done and maybe we could work together.

DR. PROFIT: One more thought about perinatal outcomes. It was like the neonatal death is so rare probably, but what you're really interested in is some kind of asphyxic event presumably, right, HIE. So brain cooling really is becoming standard of care.

You know, if you could somehow figure whether an infant gets cooled at this point, then --

DR. PRATT: Most of those get transferred out. So we probably, we would get the transfer but we wouldn't know the management.

CO-CHAIR SAKALA: Nancy had a --

DR. LOWE: Yes, thank you. Just another comment. And it came through a few minutes ago and I'm not sure you guys heard down there because of some other things that were going on.

But is the issue to think about

whether or not a concentration might be the nulliparous women at term, one baby, head down.

You know, the same one we've got the suction measure for. And because, you know, we all know very practically if we do a good job with that birth chances are the rest of her obstetrical history will be good.

I mean not always, but it's a pretty good predictor that she will labor well in the future or that she will, you know, we're more successful messing with her in causing things that we really don't want to happen than we are with a subsequent mother.

So I would suggest that that might be another place to concentrate. And it would partner well with some of the other measures that we voted to endorse.

And the second thing is, if you look on the American College of Nurse Midwives website there's a whole webpage about the Optimality Index, which comes from Europe.

But it's the idea of optimal outcomes. And what does that mean in birth. And so it is that idea of turning it upside down and saying, what we're trying to do is promote healthy moms, healthy babies, optimally.

DR. PRATT: I definitely did hear that. And actually the other thing that we may do, is then go to the other extreme and take the very high-risk. Leave the middle, the middle's muddy. And so leave the middle out.

Go to the pointy end of the stick with the high-risk and define it somehow, figure out what that looks like. Go to the high volume stuff with all those low risks and leave that messy stuff in the middle out of it.

DR. WINKLER: Okay. You've actually evaluated all the measures put forth to you. However, as we've discussed and you all have raised, we've got several measures

that are all about hospital acquired infections in newborns.

And I think the question has been raised multiple times is what do we do with so many similar measures. And this is a huge and growing issue in NQF. Our members and, frankly, it's a major issue with the board of directors.

We're really hearing and sensitive to all of the frustrations and logistical challenges out there, because we want things being measured.

But if the demands are such that we're asking to measure the same thing, or very similar things, in multiple different ways resources are wasted and people don't do it.

So there's just all the reason in the world to try and be sure that we're not trying to put out there multiple measures that are measuring the same thing differently. But rather than a library trying to help people

pick the one that seems to work the best.

So the real challenge over the last couple of years has been trying to help committees look at measures side-by-side. So I'd like you to take a look at the memo you were provided, whether electronically or on paper, on the evaluation of related or competing measures.

And what we've done is the staff has identified, in this particular set of measures that you all have looked at, four measures that are very similar.

Measure 478, which is the AHRQ measure on nosocomial bloodstream infection. 1731, which is the healthcare associated bloodstream infection in newborns from the Joint Commission.

And realize that those two were intentionally harmonized. Their differences are more about their data source and very little about their construct. So they're essentially very much the same measure.

And then the two measures from the Vermont Oxford, the late sepsis for meningitis in neonates, either for all of them, or the second measure for the very low birth weight.

So when it comes to looking at measures side-by-side, you know, how do you pick. I mean what are the principles or what are the values we use to kind of say what should get maybe more support compared to others.

And NQF has developed sort of a selection criteria. I'm not sure we've used the term "Best in Class" I'm not sure that's the right term to use. But essentially those principles are around we want to try and avoid endorsement of multiple measures unless there's a very good justification.

So that's not to say you cannot recommend, you know, several measures. However, we've got to be crystal clear as to why we're doing that, realizing the impact it's going to have out in the field for people

actually doing these measures.

NQF has a couple of preferences when it comes to looking at similar measures. There's a preference for broad target populations. Measures that capture the largest number of patients. You just get more bang for your measurement buck out of a measure that does that.

Also measures that have the broadest levels of analysis. A measure that can be applied at the facility level, the plan level, the system level, wherever. You know, has broader applicability and greater utility than a measure that's only for a clinician. Only for a hospital. Only for this.

And you end up with multiple measures that are very related but they're taking just a tiny slice. We're really looking to not have multiple measures. People are calling it measure clutter.

And so we really would like to see measures developed more robustly to be able to

capture a large segment of the healthcare system. Harmonization of related measures is absolutely essential.

And I think you can understand why. And some of the issues we've got to look at are definitional, because that's the one that kind of drives everybody nuts. You know, defining different data elements differently.

And then, of course, the corresponding codes, not the code list being off by one or two. And also because this is an evolutionary process, what we don't have that everybody wants when we do these discussions is, okay, did you run it both ways and how did it compare.

Well that hasn't been done, typically. So we're left sort of ignorant of what that really means in terms of what the results are going to look like one way versus another. But recommending that those analysis be done is a real important aspect of what we're going to do.

So, you know, again this is a very, very challenging exercise. It's a very, you know, steering committees have tried to walk through this with the four, difficult times. So I get it when you're going to struggle, I'll be there with you.

So what I tried to do is these four measures in the memo, I didn't make slides because it just got too much. I put in the side-by-side.

Now, just one other thing which always complicates it is there is another measure out in one of NQF's other projects and this is one, I think, that Craigg has referred to and that is the measure for CLABSI, a central line associated blood stream infection that is for everybody.

But it's stratified and high-risk nurseries or NICUs are one of the strata, okay. So that measure is, you know, I'm sort of close to endorsement if not absolutely. It depends on whether all the emails are in or

not.

And so that is a measure that we also have to realize is in the portfolio, so that's there. So when you look at these measures, what I did was in the first column is that measure, the CLABSI measure. And I pulled out the elements of it that relate to NICUs.

The second column are the harmonized AHRQ and Joint Commission measures. Now, we've heard from them both that they've spent a lot of time harmonizing them. There are some intrinsic differences. But, in general, they're harmonized to the extent possible.

And then again, similarly, the two measures from Vermont Oxford, 303 and 304, the constructs are essentially the same. The risk models are slightly different, because one is all of them and the other is the very low birth weight. So I left them kind of in their same columns.

And so that's how the side-by-side is laid out for you. I've presented committees with side-by-sides where the table goes on for ten pages and that really makes my eyes spin too.

So I tried to break it up to make it just a tad more readable. And if anybody's got additional suggestions on how to present these side-by-sides I'm all ears.

So if we look at just the first row we're talking about just the description of what these things measure. The first question is, how alike, how similar and how different are they? What do each of them really measure? How much overlap is there?

And then so in the first group, the CLABSI measure, you're talking about central line, or in case umbilical line associated. So if you don't have a line in you're not in the measure.

The middle group, the AHRQ, Joint Commission Measure, are blood stream

infections. And then the third is the VON measures are both blood stream and meningitis.

So there's a little bit of difference. And I'm going to ask all of you, particularly all of our neonatal colleagues, how much overlap. I mean how similar and how different are these measures?

DR. BERNS: Can I ask a question?

DR. WINKLER: Sure.

DR. BERNS: You can stop me right in my tracks if you want. But I just can't seem to get past something here. I can't figure out what happened in the past day in terms of us voting yes for 478 and then no for 1731, so if you want me to hold and we can go through this.

But my mind is still sort of stuck on that. So if folks here can help me with that I'd appreciate it.

DR. WINKLER: Yes. I think that's a discussion you need to have when you look at, particularly, how you evaluated the AHRQ

measure and the Joint Commission measure, because given that the construct's the same and they're intended to be totally aligned with some minor differences that are more related to the data source.

Actually, I've got, Suzanne just has to enter it, I actually have a table with how you voted on all of them. And we just have to put this morning's one in. And she's going to do that right now.

So if you'd like to do that first that's fine, because I do think that in order to explain your decision making I think we do have to revisit the discussion we had this morning versus the discussion on the AHRQ measure, because you did come up quite differently.

And the measures are really extremely similar in construct. Now part of this is, I think, we have to ask ourselves is were we using the same value system in our minds when we were evaluating the measures.

The conversation this morning was a little different than the one yesterday. And did we bring in new thinking that may change how we want to think about our ratings from yesterday.

So optimally it probably would have been nice to be able to discuss them, sort of temporally in the same time frame but that just didn't work out with our measure developers.

So, Scott, you know, it's fine with me if you want to try and tackle that issue first.

CO-CHAIR RILEY: But one of the remarkable things is, is if you look at, if you scan this. If you look at feasibility, seems like it may have been a pretty big driver at least for -- Do you see where I'm looking like in that second column. In 1731, Scott?

For 1731 versus 478. As well as usability so those seem to be --

DR. WINKLER: Well I think wasn't the issue with coding this morning? I mean the issue you talked about this morning was the coding and more the scientific acceptability. That seemed to be the big issue this morning.

Yet that didn't get picked up in the conversation yesterday. Did we just kind of overlook it? Or is that a true difference between the measures?

MS. KIEHN: I went back and went through my ICD-9 codes to see if we could really, truly pick up what was related in the first few days and afterwards. And that's what made me different on this one today. I don't know if anyone else had that same thought process.

As I was looking back through my notes to myself I realized that the big differences in the one we discussed today more of the aspect of could I define which ones were present on admission and which weren't

for in-born infants.

And that's what drove my thoughts and then the ah-ha that that didn't come into my thinking yesterday.

DR. WINKLER: If you had to go back to yesterday's conversation, would that impact your assessment there? Who had --

MR. GILLIAM: I think we'd have to pull it up, because I thought the issue was 731 versus 478 is definitional. It's that you can't exclude by coding those that were present on admission, even though they may not be great. But I think you can with the AHRQ. I may be wrong.

DR. WINKLER: Is the developer on the line? Patrick, are you on the line? Patrick, who was here yesterday, was supposed to call in. But the data source for the AHRQ measure is also administrative data.

MS. WATT: Hi. You know, honestly, the code tables are exactly the same for the two measures. We have worked very

closely with AHRQ, with Patrick, to do that. And actually they have always been very, very close except for one or two differences.

And we have resolved those so that there should be no differences at all. And I think, and this is sort of a self-serving comment so forgive me.

But the Joint Commission is quite often criticized, and rightly so sometimes, I guess, that we require, generally speaking chart abstraction of medical records.

And we don't do measures based on administrative data. I just think it is so ironic that we're getting gigged here because we're using coded data as opposed to requiring chart abstraction.

And frankly, I'm having a hard time bringing those two things together too. And if you could help me I'd appreciated that.

CO-CHAIR RILEY: Okay. So here my ignorance is going to show up. So if you go back to 478 from yesterday, it says in 2A1.8,

denominator exclusions. And this was the thing that I think we were all stuck on.

It says, "With principle diagnosis code of sepsis or secondary diagnosis code present on admission." So does that work or it doesn't work to exclude those, because this is what it comes down to is, and this doesn't say anything about transfer.

So what we were trying to exclude were babies that were in-born who got septic during the delivery. So does this not exclude that?

DR. LOWE: Laura, though the denominator does include all newborns and out-borns in 478. But then with those exclusions.

CO-CHAIR RILEY: It says denominator exclusions.

DR. LOWE: Right, but it does include the out-borns though. And I thought I just heard you say it was only the in-borns.

CO-CHAIR RILEY: These guys, when

we were trying to figure out how to make the JCo one work, we kept getting stuck on this sort of transfer thing. And the transfer works, because that made sense.

So I'm just wondering if, the way this is written, the denominator exclusions exclude cases with a principal diagnosis code of sepsis. Or a secondary diagnosis code present on admission.

DR. ARMSTRONG: I think Jaleel led this discussion and he said that when the baby is admitted to the NICU that present on admission is not yet confirmed. It's suspected but not confirmed, so it's not going to come in with the code. And hence you can't --

CO-CHAIR RILEY: So then this should look the same as the other one the. So this has the same fault or same flaw, I shouldn't say fault, flaw as the other one. Is that right?

DR. GREGORY: But there is a teeny

bit of a distinction and it would be great if we had a coder here to confirm I'm telling you the truth.

But a newborn that is the principle diagnosis. A baby born in your hospital, a newborn, is the principle diagnosis.

Then they could be admitted from labor and delivery with sepsis, okay. That would be excluded from this. New conversation. A baby transferred with sepsis has present on admission.

And it's present on admission is new within the last, actually ever since we started coding for DVT and all these diagnosis that are hospital acquired, they've been extremely vigilant about applying the present on admission diagnosis to get those exclusions.

So the present on admission is to capture your transfers and to exclude them. And then the principle diagnosis for newborn

baby is newborn baby. And sepsis would be a diagnosis that the neonatologist would have had to given that baby from labor and delivery.

FEMALE PARTICIPANT: So if that's the case then that flaw is the same.

DR. BAILIT: Well no, because part of the problem is, and tell me if I'm telling tales out of school, is that you can have a newborn baby that looks fine on admission and two days later becomes septic and that's still an early neonatal sepsis.

And that would not be in your present on admission or even in the admission codes. But it still counts against the OBs and not the pediatricians.

DR. JALEEL: I agree. I think what we want to have a dichotomus is early onset sepsis versus late onset sepsis. That's what we want to know.

And what struck me when we did this discussion today, in the morning, was the

title of the Joint Commission measure, which said, "Health Related Blood Stream Infection" or something like that, I don't have it up here. Healthcare Associated, which means it should not include early onset infections.

And over here the 478, which we did not, probably it was one of our, we all looked at and partly the reason it might be because it says neonatal blood stream infections and not health care associated.

MS. PARTRIDGE: It sounds as though we're revisiting our decision on 478. Is that an option? I mean we might say no to both of them?

DR. WINKLER: I think that it's going to be very hard for me to explain to the audiences out there for two measures that are as close as these are to have two different decisions. So I would like you to reconcile it. I think that's an important conversation for you to have right now.

You know, whichever way you go.

But if they're not the same it needs to be based on the differences of the measures. And they are very much alike.

DR. PROFIT: I think what struck me today -- I'm sorry. Just one second. So first I think I really want to recognize that both the Joint Commission and AHRQ I think have done a really excellent job at putting this together and risk adjustment and getting expert input and all this.

I think it's a really thorough process. What struck me today that didn't really strike me as much as yesterday and probably led me to at least mention it, even though I didn't change my final vote, was that I was concerned that there is really no- sort of, and we talked about this a little bit- like criteria on validation being done.

So we know, we have coding that is probably is done as well as you can do it. But have we really compared those against a gold standard? Or at least tested it.

So my hesitation is a little bit about making this a measure for public reporting and comparing hospitals when we just have very little information about how much this actually reflects the truth as close as can get it at least.

And so that would be my concern with regard to putting this out to the wider community and that we'd want to see some kind of information like that.

FEMALE PARTICIPANT: Reva, we missed the public comment period. And I know you have a very rich discussion going, but I had comment that's related to this discussion. I am here representing the Association of Women's Health, Obstetric and Neonatal Nurses, which it represents the interest of over 350,000 nurses, as you all know.

And I'm very concerned that we have the Joint Commission measure has been well adopted, a lot of hospitals may not have been reporting it yet publicly, but they are

working on it.

And very concerned not to have that measure go away. And also, especially since they are so similar, it's so similar to the other measure unless there was some extremely good reason not to continue to continue to measure this, I think, it would be important to have this for the country.

The issue of the patient's having infection or the locus of the infection I think would probably be more compelling to me if the rates of these infections were higher than they are.

So the total rate of these infections for the hospitals, the 164 hospitals who have publicly reported, is only 0.3 percent. So I think we might be nitpicking here a little on something that may not be having a huge impact.

I don't know, but that would be something I think for future discussion, for getting other data on that. And without the

measure I don't know if we'll be able to get those data.

MS. WATT: You know, this is something that we have talked about with AHRQ, believe it or not, in terms of identifying these patients. Whether or not we should be adopting the present on admission, which is not generally in the medical record. It's a billing designation required by CMS.

And that's part of our issue, because our data source is the medical record, it's not the bill. And that is the big difference actually between this, our, the Joint Commission measure and the AHRQ measure.

But based on the discussion I hear you, we hear you, in terms of where you're coming from. And I think that what we can do -- excuse me, is to evaluate.

And actually we sort of have a golden opportunity to do it right now, because as you know we are in the middle of the reliability test of this measure.

We can actually look at patients in the next, I think we have six more hospitals to go to, you know, we can look at those patients who meet the numerator for this measure and look to see what proportion of that group had positive cultures within the first three days of birth, or whatever, to see if that is a big proportion.

And if, therefore, skewing the measure rates. You know, we can do that, we're willing to do that. And we can certainly, of course I'm saying this like yes, it's like that.

But we could, I suppose, add a data element that would have to be manually chosen, picked out of the medical record that says this is a intrauterine infection.

Now that would be pretty hard, I think, to define that data element for a non-clinical data abstractor. But, I mean, we're certainly anxious to make this the best measure that it can be.

DR. PROFIT: So I think my general concern is even lesser about those first three days, because I think it would be important to look at what you said during the reliability testing.

But I guess my personal feeling, which I don't know would reflect the wider general population, would be that those numbers are probably relatively small overall.

But then I think what I would be a little more concerned on is like later on. How reliable is an identification of an infection in the NICU. How correlated is that, like you know, whatever.

Maybe CDC or other places have defined what they think is sort of the gold standard where whether something is truly and infection, like -- things like positive blood cultures.

You should have a couple of algorithms to define whether or not, but I think as you said, you could do this now if

you did reliability testing. You could look at the charts and determine against a gold standard how valid this is.

DR. JALEEL: I think that's a good point. But there are two components to the early onset septicemia. So one is these are vulnerable babies who are on extremely critical nursing care. And stopping antibiotics on these babies is something which gives jitters to everybody.

So many of these babies are continued on antibiotics just because they are sick. They don't want to, a provider, does not want to stop the antibiotics. So that will also be coded. If it is coded it is coded as neonatal sepsis.

So that is one group. And the second group is a group which is positive blood cultures. So that's confirmed early onset sepsis.

So if we can get both of these together, and that's a big group when you say

the number of babies who just receive antibiotics for presumed sepsis, that's a big number of babies.

So if you can identify both of those from the reliability testing that will be good.

MS. WATT: You know, maybe this is sort of a moot discussion anyway, because let's face it, by 2015 they say, or even if by 2018 or whatever, these measures are all going to be specified for retrieval from the electronic health record.

And what we can't do now, with the existing ICD-9 and 10 codes, in terms of identifying positive blood cultures those kinds of things, we will be able to do with the electronic specifications because that will be a LOINC, you know it will be defined in LOINC.

And so I don't know, maybe this measure has been in place now for, well a year. And it seems to have good resonance

with the country. And we know that it's going to be transitioned into an electronic measure sometime in the foreseeable future.

Is it, and I'm asking you this, this is not a rhetorical question. Well this is what we say at the Joint all the time. Is the juice worth the squeeze to try and do this at this point?

DR. GEE: I just wanted to add from a state perspective that there is such a dearth of measures that can be arrived at without direct chart abstraction. And we are hungry for those types of measures.

Those are the kinds of things we're hoping to put on report cards to drive change faster. And I don't want the perfect to be the enemy of the good here where we're not, I applaud the efforts that you have made to make this ICD-9ish.

And so I just wanted to put that out there. We need more measures like that.

DR. BERNS: I'm trying to go

through these four measures again in my mind, it's getting jumbled here. But the criticisms, the points that you guys are making today, I'm looking at the neonatologists here, you know, I'm a pediatrician.

So don't they apply to 303 and 304 too? I mean, how's that --

(Off microphone discussion.)

DR. BERNS: But isn't this based on. I mean help me here, I'm really being dense. So this is based on septicemia or bacteremia as well. But it's --

DR. GEE: Rule out sepsis, right? Isn't that what this is all about? Put the baby on antibiotics just in case it's septic.

DR. JALEEL: This is coded as neonatal sepsis. But without a positive culture. So if you're 303 and 4 that is with a positive culture. Either from the mother or from the --

DR. BERNS: So it's suspected

bacteremia is that what?

MS. WATT: I think just to talk about it in developerese, I think the issue is really you're looking at can we trust that what the physician writes in the medical record as an infection, and therefore is coded as an infection, because I don't think coders just code stuff willy nilly, is that right.

And you know what? That's really not what this measure is intended to look at. I understand what you're saying and that there might be a problem here. But I think that as far as these measures themselves are concerned, they just reflect what the documentation is in the medical record.

And they don't intend or try to identify whether or not that documentation is correct. I think you're making, I think you're not looking at the measure as a measure itself.

MR. GILLIAM: Just for clarification, and I do want to ask one

question. With VON it's clinical people that are doing the data collection if I understand that correctly. It's not --

FEMALE PARTICIPANT: No.

MR. GILLIAM: Not necessarily? I know in my facility it's --

DR. JALEEL: When you say clinical?

MR. GILLIAM: I mean RNs that are collecting it.

DR. JALEEL: In our institution it a research nurse, but I don't know in all other places what the code personnel.

MR. GILLIAM: But then my question is, is 1731 and, as far as that goes, 478 are those blood stream infections that you're asking about. Are they blood stream infections related to the hospitalization? Because VON, in my opinion, is looking at those related to care in the hospital.

Something that we can do something about and improve performance. With, as I

said, maybe 900 facilities, even though it's a subset of facilities. Whereas this, I'm not sure, it's just looking at the global blood stream infection, if I understand it.

DR. MURI: No, that's not entirely true. What we are attempting to do, what AHRQ is attempting to do, is to eliminate those patients who are admitted to our hospital already infected, whether that means they are coming out of the uterus already infected or they're coming from another hospital already infected.

And the problem is there is no way to differentiate. There is people who are transferred in and I think both of these measures do a good job of that. And I think what the concern has been this morning is that there is not a good way of identifying patients who are infected intrauterinely. And VON can't do it either.

CO-CHAIR RILEY: But actually, I mean, I think --

DR. HORBAR: Can I just say that that's not quite accurate. We collect information about early sepsis separately from late infections.

DR. JALEEL: I would like to mention two things about the measure title of the Joint Commission measure. One is it healthcare associated. If you are deriving this information from coding then it is a combination of both early onset sepsis and late onset sepsis. And this presumed sepsis.

So that is not clearly, purely healthcare associated. And the second thing is blood stream. If you're deriving this from coding it is not a blood stream infection because it is not a documented blood stream infection.

The code just says neonatal sepsis. So it's not a blood stream infection either. So the measure title itself is not accurate.

MS. WATT: Well we can certainly

change the title of the measure.

DR. PROFIT: I don't think it's just the title of the measure, I wouldn't be so nit pick, personally I don't feel that nit picky about that.

But I feel for national quality benchmarking, public benchmarking, I don't really feel like I'm being extravagant in asking that coded data be validated against like a gold standard.

You know, especially if you have the opportunity to do it, like right with the charge that you're reviewing right now.

CO-CHAIR RILEY: I have a question about that.

DR. PROFIT: Go ahead.

CO-CHAIR RILEY: Because initially what got us off to sort of a different tangent on 1731 was the sense that what you would try and change, if you had a high rate at your hospital, what you would try to change would be whatever practice is happening in your NICU

that's giving your babies infections.

And so what you wanted to do was carve out those babies who actually got infected in the NICU and not include the babies who got infected in the uterus. So that's how we started the conversation down this path.

It's not clear to me how the validation is going to get to that. And I actually don't even think you can get there, because if you think about it clinically we can't figure out who's even infected in utero.

If the mother's even infected, or the baby's infected and then the baby comes out and it takes, you know, you'll never be able to sort of even clinically answer the question who got infected in the uterus and who was really the baby who happened to be there and on day four the kid's sick.

Most of those kids are on antibiotics anyway, right? Just because they're pronged or because they're pre-term

labor everybody says they must have been infected.

So they're already on antibiotics. So I'm not sure that the issues with the measure --

DR. PROFIT: I think we're talking about two different issues. So I think there's two different issues. I think there's one, there's the early onset differentiation for treatment of suspected sepsis during the first maybe 48/72 hours of life.

And then there is an issue, to me, about the data validity much later on during the course of the hospitalization, where you have, and maybe you can educated me how the coders usually handle that data. You know, we have a resident or a nurse practitioner, attending, multiple people charting on the same patient.

If the baby has rule out NEC, you know, the baby got started on antibiotics. Somebody writes, like, rule out sepsis at the

same time because we don't know if the baby has NEC or sepsis, you know, what it is.

Low bowel sounds, so it gets started for like a 48-hour watch. Not really sure, you know, Stage 1 NEC is essentially defined as a hodgepodge of it could be almost anything.

And then the attending may not feel like that this baby really has an infection but somebody else has coded it as an infection. The cultures are negative, you get five days, six days, seven days of antibiotics and it may permutate to the coder extracting it as an infection when the baby never really had a positive culture or an infection.

And so I just feel like there may be issues along extracting, typical for coder data,, I guess, just extracting whatever the truth is from the chart.

CO-CHAIR RILEY: So are you saying that if that part of the study was done, if they went back and validated the fact that the

coding, wherever they got it from, actually reflected some positive blood cultures that you would be fine with it?

DR. PROFIT: Yes. Yes and I think I would need --

CO-CHAIR RILEY: How, then, do you answer the first part of the issue. Is that a big deal or no?

DR. PROFIT: I mean, that's what I've tried to say before, is that I think the actual incidence of early onset sepsis is probably very low. I think Joe Carpenter is on the phone from VON and he may be able to give us numbers.

I'm not sure what -- but the times we truly see positive blood cultures at birth are probably not that terribly high. So I don't know, compared to the overall infection rates, whether that makes much of a difference.

So my real concern is mostly about just the validation of the -- and I think

Jaleel has more concerns about the early onset infection maybe than I do. And I understand his reasoning.

So I feel like -- I think these are two separate questions, slightly, but I think both of these questions could be answered during the chart review.

DR. SUTHERLAND: But if you don't mind, I just want to bring up the issue about UHC. And I don't know if anybody here participates in UHC. But the problem is that what we're seeing within institutions is using administrative data, because we may have coders who are very talented and they've been told to capture every single piece of information and code it.

For example, accidental punctures and lacerations. I don't know how many of you are following that. But we're finding in our institution is doing a laparoscopy and "rupturing an ovarian cyst" is coded as an accidental puncture/laceration. Or cutting

the uterus when you're actually taking it out. So we're doing a big documentation project at the Cleveland Clinic to try to have communication with our coders to understand what's a technical error that really needs to be coded as an accidental puncture or laceration versus something that's a normal course.

For example, the normal course in a very pre-term baby is to put them on antibiotics. That would be read out, though, by your coder as probably a sepsis diagnosis. So I think that's really the concern here. I think when you get down to it our coding data doesn't really reflect on the truth.

CO-CHAIR RILEY: I mean our you saying we don't whether or not the coding data reflects --

DR. SUTHERLAND: Right, until you really get to the level of the chart review.

CO-CHAIR RILEY: So then the only thing I would bring up --

DR. SUTHERLAND: -- many of them fall out.

CO-CHAIR RILEY: -- is that we have the same issue then, 478 and 1731, suffer from the same -- if that's the real concern, both measures suffer from the same issue.

So we probably need to go one way or the -- I mean, it makes no sense to have -- And the title is not the reason. Do you know what I'm saying? So I think if we --

DR. JALEEL: That was -- I just mentioned about the title because that was the prompt for me. But, yes, I agree. If the coding is the same in both these places, then I think both need to be weighed the same or valued the same.

CO-CHAIR RILEY: Can I just ask --

DR. ROMANO: Hello? Can I speak?

CO-CHAIR RILEY: Yes. We're listening.

DR. ROMANO: Oh, hi, this is Patrick Romano. I'm not sure -- I just joined

about 20 minutes ago. I'm on the phone from California today, again representing AHRQ. Can people hear me?

CO-CHAIR RILEY: Yes, we can hear you.

DR. ROMANO: Oh, thank you. Well, I just wanted to point out one thing that I think may not be apparent in this discussion, which is that the logic of this indicator, in both of our harmonized specifications, requires a specific organism to be identified.

So I don't think the concern about culture negative infections applies. If you look at the architecture of the numerator specification, it requires the identification of staphylococcus, either methicillin-resistant or methicillin-sensitive, a gram negative pathogen, such as E. coli pseudomonas or serratia, or disseminated candidiasis. So there's no way that this could be triggered unless there's a code for a specific pathogen.

And, of course, a coder would have no basis for applying such a code without a positive culture result. Obviously, we look forward to the opportunity to validate this in collaboration with the Joint Commission.

But I don't think that people would be sort of inventing organisms as they code the record. Logic does require that if the hospital reports the diagnosis of neonatal sepsis they must also report a separate code documenting the pathogen that was identified. Does that make sense?

CO-CHAIR RILEY: Yes.

DR. PROFIT: That's very helpful, thank you.

CO-CHAIR RILEY: That is extremely helpful and answers your question, does it not?

DR. JALEEL: Is that true for the Joint Commission measure as well?

MS. WATT: We have the same codes.

CO-CHAIR RILEY: It's the same.

DR. ROMANO: We've harmonized ICD-9s specifications. The only difference, as Celeste has mentioned, is that in the reporting to the Joint Commission the hospital would have the opportunity to override what's in the administrative data if it felt that there was an error. Am I saying that correctly, Celeste?

MS. WATT: This is Ann, Patrick. And actually not entirely. But generally speaking we don't give abstractors the capability to say, oh no, that's not the right code, therefore I'm not going to use it.

The intent of our measure is that the codes, as they exist, are the codes that are used.

DR. JALEEL: I think that was -- I think that was not clear when we discussed this with the Joint Commission measure, that the concern I had, which I clearly voiced out, was that there will be a significant number of babies who will be culture negative and will

be coded as neonatal sepsis. So it was not clear when we did the Joint Commission measure.

CO-CHAIR RILEY: So can we re-vote?

MR. GILLIAM: Can I ask one question? Since you said they were the same, in the document that was submitted where is that? I found it clearly in 478, but I can't find it in 1731.

MS. WATT: Where is what?

MR. GILLIAM: That -- as Patrick was saying that you have --

MS. WATT: It's that -- excuse me for interrupting. It's those tables that we reeled off this morning. They are exactly what Patrick just reeled off. So 11-10 to 11-11.

MS. KIEHN: Yes, 11-11 is the one, exact same thing.

MS. WATT: Streptococcus Group D, enterococcus, staphylococcus unspecified,

staphylococcus aureus, other staphylococcus, Friedlander's bacillus, klebsiella pneumoniae, E. coli and pseudomonas. They're all there.

MR. GILLIAM: But it's in a separate document, it's not in this submission?

MS. WATT: Yes, it is in this submission. It is in a table that is referred to in this submission and was included.

DR. ROMANO: And I might hope to clarify that what we tried to do with our expert panels, and I think the Joint Commission had this discussion as well, was to specifically exclude some infections, such as Group B strep or listeria, that are overwhelmingly perinatally acquired.

Now, of course, thanks to antibiotic treatment, we don't see nearly as much Group B strep as we did when I was in training. But, still, those organisms that are specifically perinatally acquired are excluded from this list.

DR. BERNS: Yes, 478 had the list in it. Whereas this was -- you had a link to it to get online to the table.

CO-CHAIR RILEY: Can we go back and now vote on 1731? Are people okay? Have we discussed this to death now?

We go through the whole thing right?

(Off-microphone discussion.)

CO-CHAIR RILEY: So this one we all loved. Let's go on to the next ones. Scientific acceptability. Right, because that's where it started to go in the wrong direction. Okay. So now we're re-voting on 1731 for Scientific Acceptability. Yes.

Everybody press your button again. I don't even know how many people are still here though. We are 21? Okay. Perfect. So it's 21 yes and no noes.

Can we move on? And now we'll do Usability. 9 high, 12 moderate.

And then we'll go on to

Feasibility. We're there, 7 high, 14 moderate.

And now overall suitability for endorsement. Here we go, 21 yes, no noes. So then it looks a little bit closer. Okay.

MS. PARTRIDGE: So we've now voted yes on both? I'm sorry, on all four. And we're now trying to narrow that?

DR. WINKLER: Well, you've recommended four measures, which you have identified as being similar. And you've already brought up the issue that that can cause issues out in the field. Implementation, you know, demands on hospitals.

I think you need to have a further discussion about the implications and the impact of the recommendations you're making on these very similar measures.

I think one of the first questions I wanted to know, because I really am not totally clear, is how much overlap? Are these really that similar? Is there something

really different about some of them, maybe?

Because it really is the 478, 731 we've already just determined they're the same. And 304 and 303 are pretty much the same. It's just one's a subset of the other for very low birth weight.

So the constructs are pretty much the same. The question is are they measuring the same thing? Or are they measuring something different?

DR. SUTHERLAND: And before we get too far, how does CLABSI factor into this?

DR. WINKLER: Well, realize that there is a measure also that includes NICUs for CLABSI. So that's just another one that's out there.

You can't act on it, but to know that it exists in NQF's endorsed measure portfolio is -- again, is related.

DR. DENK: My understanding was, from yesterday's conversation, was that one of the things that was unique about VON is that

it's sort of run more like a registry than anything else.

And that there's additional criteria for patients in the VON network to sort of be sort of counted towards all kinds of data collection.

So if we asked if that's true, and we ask VON to harmonize with these others, then they're going to de-harmonize with a lot of other things within the VON measurement system. And that seems to me to be a reasonable excuse to let it be slightly different.

DR. GEE: Another difference is VON has a number of modifications for gram negative bacteria, which we've mentioned are very common. Often it's unclear as to the severity of those infections.

And so VON has spent more time thinking about how risky are those and which ones are meaningful versus this measure which doesn't. So that is -- you're measuring two

slightly different things.

DR. PROFIT: So I would say I think the CLABSI measure is a little bit different. I mean, maybe you could argue that the underlying processes are still rather similar, but CLABSI is used like in large state collaboratives. Now there are specific bundles for it. So I think because it's restricted just to line infections it's a little bit different.

I think the other measures -- I would say, a large degree of overlap overall between the other measures, personally.

I do. I mean there might be slight differences. But overall I think they're all measuring blood stream infection. A pretty large number of at-risk babies.

I don't think I can like easily say, based on the scientific merits of these measures, whether any would be more superior than the other.

I think it kind of comes -- to me

personally, it would come down to issues such as feasibility and what's already being done.

So the easier it is to get or the larger number of places already doing some collection would, to me, trump my ordering in my head for this. But I don't think we need all four measures for all four hospitals.

DR. BERNS: Yes, I think, Chuck, you actually had a question for the VON folks, right?

DR. DENK: Yes. Okay. So that should have been in the form of a question, yes. Is there definitional issues about how patients are counted in the VON network and excluded from both numerators and denominators as general things that would make it hard to harmonize with these other two measures?

DR. HORBAR: Well, I'm not exactly sure how to answer that question when you say are there definitional issues. I mean, are you asking what are the criteria of eligibility for the VON?

DR. DENK: Yes, I think that's kind of what I mean.

DR. HORBAR: Well, basically it's any infant that's born in the hospital that's within 401 to 1,500 grams or 22 to 29 weeks, or who's admitted to that hospital within 28 days of life. That's the very low birth weight measure.

We've reported for all the very low birth weight infants, as well as for infants 501 to 1,500 grams, for example.

Now the other measure, 303, includes any infant that's admitted to a NICU within 28 days of life at the reporting hospital. Or who dies at the hospital on or before day 28. As well as the very low birth weight population I mentioned.

DR. WINKLER: Now, one of the things I think is difficult about this whole issue is the fact that all of these measures are -- they're slightly different data platforms, which means they're likely to be

somewhat different users.

The AHRQ measures are clearly used often the state level and, I don't know, purchasers use them. And so, because they're purely administrative data based -- Rebecca mentioned that.

The Joint Commission's are administrative plus. So there's some, perhaps, some extra chart review in using their specific methodology. But we've already determined that the two of them are aligned as much as possible.

The VON measures are just completely different data source. And, you know, it's a registry. We don't have any head-to-head comparisons. I guess, you know, does that mean they have different users?

DR. PROFIT: I don't mean to speak for, I would ask VON to respond. You're looking at me but --

DR. WINKLER: I'm looking at you because your hospital is -- how many of these

does your hospital do?

DR. PROFIT: I assume we do all of them currently, probably. I mean, I'm not in the measurement department of my hospital. But I know we do the CLABSI ones that go to the CDC for sure. You know, we code all these things so I'm sure AHRQ could -- they take, I'm sure, probably part of the Joint Commission also.

So we're probably doing that effort as well. But I'm just -- like, I guess as a value -- what I'm trying to think about is as a value to the frontline provider who's trying to reduce healthcare infections, is like I don't need four of these.

That's like -- maybe CLABSI plus one of the others. You know, I think like the VLBW measures, maybe -- like before I said they all can overlap. But maybe like there's a VLBW measure from VONs a little different because all the other measures cover larger baby also.

But otherwise, you know, I think the general gist of this is that the other users, like the VON, are largely going to be the NICUs directly. And maybe in extension to the NICUs the hospitals that ask, you know, the board of the hospitals to ask the NICU director to present that data.

And then through VON there's also families in a lot of NICUs have access to that data. But that's dependent on the NICUs' use.

DR. SUTHERLAND: I guess my issue is are we really targeting the line use, because when we look at the part 1A-3 on 1731 there are five citations. And it says that the risk factors are decreasing birth weight, central venous catheter use, prolonged perineal nutrition and prolonged ventilation.

So is the point really to target careful use of central lines and reducing central line infection? I guess that's really my question, even to Joint Commission.

DR. PROFIT: Yes, I think we know

that the overall infection rates, if you look at just CLABSI rates versus overall infection rates, there are substantial differences. And I think CLABSI is also now being used as part of U.S. News World and Health Report. It is, right?

So we haven't talked about CLABSI but when you talk to IUD specialists within each hospital there's some leeway to decide on what are we going to call a CLABSI, what are not going to call the CLABSI. So there's going to be some wiggle room around those things where you may not have perfect data from each hospital either.

So it comes back, I think, to the philosophical idea about what are we trying to do. And I don't know if I have a perfect solution to this.

I think all of these measures are reasonable measures and they're well -- you know, obviously all of them have a lot of input from a lot of people and they're well

developed.

I would just, if I was a hospital CEO, I'd want to go ahead and shoot myself over having to report four different measures.

DR. WINKLER: So how about Teri or Jenny, you guys --

DR. PROFIT: Maybe not shoot myself, but --

DR. WINKLER: -- are often on the front line for this stuff.

MS. KIEHN: I know that we do provide the CDC measures and we do do VON also. So having to do -- we have chosen not to do the Joint Commission just for that fact that we know we have to do the CDC.

VON we use internally for data, because we're already collecting a lot of other measures with it. So that's my take on that side.

DR. JALEEL: So I agree with Dr. Profit, the CLABSI is sort of kind of different measure. And it would probably be

useful to have that as a separate measure.

But these four, there is a significant amount of overlap, if you have a Venn diagram, most of them will be covered by all of these four measures.

So then it comes to me as to if I had to pick one, which one would I pick? If I just had one choice. Then I would say 304, which is late sepsis and meningitis in very low birth weight infants. So that's my preference, but again --

DR. GEE: You also have to think about real world application. I would argue five years from now we don't need different ones, but now we do. In Louisiana, I just paid to aggregate all of VON data. We have 17 NICUs who report. It's an excellent measure.

But for the half of our NICUs it's just not approachable, we can't get -- so are we going to make a decision that will exclude other metrics that use ICD-9 or -10 coding? Because VON is slightly better?

And then make a decision that we're not going to get data from our hospitals that don't participate in VON, which is associated with a cost.

Or do we just say right now, in our real world healthcare measurement world, we have different ways of collecting data. And also, hospitals are not required to report JCo measures now. They ought to be, I think, on a lot of these required, but they're not now.

And so if they want to make a decision to report VON they'll do that. If they don't report VON they'll do what's convenient for them. But at least it will give us some data that we can use for quality improvement.

DR. JALEEL: I completely agree with you.

DR. DRYE: Yes, I do too. And I would just add this comes up in the measures that my group works on as well, because they

are registry measures and they're claims based measures.

And sometimes actually we're intentionally developing them in parallel because usually only a certain portion of hospitals are going to participate in registries.

We don't want to undermine what they're doing because sometimes, and oftentimes, those are the very best measures. But we don't wait or require every hospital to join a private registry.

So I think the marginal costs to the hospital -- I think it's difficult to look at four different -- it's really three different measures results because the AHRQ and the JC measure are the same.

But I would just say right now I'm not hearing that it's a lot of a marginal burden. If you're in VON then you get this measure result. And then the other ones are claims-based, so it's not a burden on the

hospital.

So I don't think we're generating an excessive burden by allowing all these measures to stand at this point.

MS. BRANDENBURG: I think the Level 3 centers in our area -- we've already gone to VON so that's what they would choose. But the Level 2 centers they're most likely going to choose AHRQ or Joint Commission, so it just depends on the facility.

DR. JALEEL: Yes, and that was the reason for approving both yesterday. That was the discussion that we had.

CO-CHAIR RILEY: So it sounds like everybody's saying we need at least two, right? So at least 303, 304 and one of 478 and 1731, am I hearing that reasonably?

MR. GILLIAM: Just to confuse it slightly. As Reva said, CLABSIs are required in many states. And in fact, beginning first of the year most hospitals in the United States will have to collect CLABSI data, and

whether their state requires it or not for CMS they're going to have to collect that data.

And most pediatric hospitals do VON, or they're in VON and so they're also doing the CLABSI because they're required. And so we're in the same situation, we chose not to do AHRQ or the Joint Commission because we're doing the NHSN system.

And it's separate groups, so even though it's a third measure -- and my concern with all of this is you will get slightly different rates or percentages of infections depending on which one.

VON, which is very good, they have definitional problems that the CDC does not like, as far as goes with staph epi. AHRQ and the Joint Commission, using coding data, you only have to have one positive culture. In NHSN, depending on the pathogen, you have to have more than one positive.

So whatever is recommended, it depends upon whose ox is going to get gored,

so to speak, when it's presented at a board, because your numbers may look good with one system and not so good with another.

CO-CHAIR RILEY: So at the end of the day, we can't really control what people do with the data, who gets gored. But we can at least make sure that the measures are available so that they at least have a menu and they can choose. Would that be fair?

MS. PARTRIDGE: I don't want to lose Rebecca's comment, because I'm sitting here wearing my old purchaser hat. I don't have access to VON data if I'm a state Medicaid administrator. I don't necessarily have access to J Commission data, but I'm more likely to.

DR. GEE: And with VON it's totally -- it's blinded. So even though I paid for the data I can't report which hospital it is. That is the mantra of VON.

And so for a public reporting -- and it's fine. That's good for quality

improvement, but for public reporting it's useless. Unless -- I mean, you can see trends, but you're not going to be able to see which hospital can I go to for better care?

MS. PARTRIDGE: What I wanted to say is as this goes forward, if we end up recommending all three measures, I think the record should be very clear that there are considerations for hospitals, which are entirely reasonable, you know, I'm not going to do Joint Commission because I'm doing VON.

But at the same time what we're hoping is that this is accessible to purchasers, public and private. And therefore we don't want to discourage the adoption of a very similar measure that could be publicly available, such as the Commission's.

DR. WINKLER: I guess the question I would put to you at this point is, do you feel strongly enough that you really need to make a selection among the four? Can you live with the push back you're going to get by

recommending four?

DR. GREGORY: This is Kim. Am I really recommending four or I've endorsed four, I mean, they can do whatever they want.

DR. WINKLER: Yes, right. And that's a philosophical issue around NQF, it's been there since the very beginning. Do we create a library or are we trying to be more directive.

And that comes and goes. But I can tell you that -- to standardize so that everybody's doing the same thing. It fosters comparability, fosters standardization. So this is a tension that's very real.

MS. PARTRIDGE: Right, and that's why I said what I said about making sure the record is clear as to why, if we go forward with four, we are going forward with four or are we going forward with three, because as it goes up to CSAC and as it goes to the Board there will be questions raised as to why all four or all three.

DR. WINKLER: It's not that these issues aren't valid. I mean, with justification. Now not everybody is going to agree with your justification.

But as long as we can clearly articulate the reasoning, and that's what you've done over the last hour. And so I think we're there. I just need it to be your decision.

DR. PROFIT: I have a question.

CO-CHAIR RILEY: Is there any way to choose between 303 and 304? No?

DR. PROFIT: It depends on what you want to measure.

DR. SUTHERLAND: I guess my concern with leaving all four on the table is you're leaving the opportunity to game the system. Okay?

And I know that sounds very negative, but I'm in a situation where we're bringing together a lot of hospitals who have different backgrounds as far as how they track

things.

And they may choose, they say we've always done this, therefore we do this. And we're going to look unfavorable if we switch to that. So this isn't something that I directly am involved in.

But when we talk about, even within internal hospital systems, you know, leaving so many options out there I think makes it harder to align care and to really compare.

You know, we really should be comparing apples to apples. I guess that would be my concern.

DR. BERNS: Yes, I think I get what you're saying, Sharon. I don't know if I totally agree. What I'm struggling with are the criteria that you put out here for us, Reva, which I'm specifically thinking about how to reach as many possible individuals and entities.

And if that's what we really want to do then I feel sort of stuck. Because I

really want one. I really want one. I want one.

But you have however many number of hospitals, 900 or whatever it is, that are already reporting into VON. That's huge. I mean, you know, that is just huge.

At the same time that's not public data. Right? So where's the accountability piece, it's up to the hospital. So that's where I get torn, because it's like -- I would like one.

I'm thinking, okay, do we make it just one topic and then you could have a couple of choices underneath? That's going to upset people too I'm sure, or confuse people.

So that's why, I mean, the fact that we're harmonized with AHRQ and Joint Commission, I think, of that, Laura, as one just like you. And I would love to see a choice between 303 and 304 if we could. I hear that that's not as helpful from a practical standpoint. But we are going to get

push back if we approve three.

And we do have to decide, Reva, if we can live with it. And, for me, I think I can live with it with the rationale I'd like to as well, but with the rationale of we're reaching as many possible individuals and entities as possible. But it's that public reporting piece that we just don't have through VON, which is -- yes.

DR. GEE: My understanding was it's not four measures. We're talking about three that we then harmonized, right, Patrick? And that has already happened. So when you say four, let's at least make it it's only three.

DR. WINKLER: Unfortunately, that doesn't fly, because two is two. They need to be harmonized, but for those folks who count the 700 measures in our portfolio and wonder why in the world we have so many, it's two.

DR. ROMANO: This is Patrick. I mean, I think -- I don't know if the 700

number is some automatically evil number. But the point I think is to have measures that can be applied by a variety of different users and stakeholders.

And what we tried to do is set things up so that if a hospital chooses to participate in the Joint Commission perinatal measure set, they have a measure that they can use.

If they chose not to do that, if an insurance company or a researcher or some other organization that has access to administrative data, all payer administrative data, wishes to look at the same phenomenon, they have a measure that they can use. And those measures have been harmonized.

So I don't -- in a world where we have different data streams that are under the control of different organizations following different paths, I don't quite see a way to get around this problem. At least until we reach the nirvana of the electronic health

record where all of this information can be accessed by anyone.

DR. GREGORY: The point that keeps coming up about VON not being publicly disclosed, isn't the whole point of them having brought it here is so that it ultimately would be publicly disclosed?

DR. WINKLER: A couple of avenues. A measure that's endorsed by NQF, we make those measure specifications available. They could be adopted by non-VON folks to be implemented in their facility and used and adopted in some other way.

So we aren't endorsing the VON registry per se. And then each of those potential users determines, you know, how they will use in public reporting.

Now, certainly one of the things we encourage, and we're looking for measures to become more publicly available, and that would be one of the things we really want to encourage. But it's not absolutely required.

DR. PROFIT: So that was kind of my question. So if Medicaid, you know, so VON is proprietary but if Medicaid said, well, we want from all hospitals data specified just as the VON does, or like the Joint Commission does. What keeps them from theoretically doing that?

Is it really our role to say we want to measure for X, Y and Z so that every stakeholder can get at it? Or is it a role for us to say, okay, we think X, Y, this does seems to be the measure that makes the most sense, as it currently is.

Now it is up to the policy makers to take it further. You know, I guess I'm not quite clear where the overlap or the frontier there is.

MS. PARTRIDGE: Unfortunately, Jochen, it's not quite that simple. If I said I want all the hospitals to use the VON specs and submit the following data to me, I wouldn't have any way to gather it, aggregate

it, sort it, analyze it or spit it back out.

It's not just the specs that you're buying in essence. What I would like to have is: let there be an agreement between VON and X that you'll send the data to VON then VON will share it with whomever I, the hospital, have agreed can have it.

That would be the nice way to do it. But presently that's not the way it usually works.

DR. JALEEL: Can I make a suggestion? So I like the 304 measure and VR part of VON as well, and we appreciate the work they are doing. But for simplicity's sake, can we do this that, since all these four measures almost measure the same, harmonize 478 and 173 and pick that? What's the downside of doing that?

CO-CHAIR RILEY: So you're saying what Scott and I were sort of motioning to, right?

DR. JALEEL: No, just one.

CO-CHAIR RILEY: Oh, just one.

DR. JALEEL: Harmonize 478 and 1731, which they are.

CO-CHAIR RILEY: So that's one.

DR. JALEEL: So that's one.

CO-CHAIR RILEY: And so just pick that?

DR. JALEEL: And pick that.

CO-CHAIR RILEY: To select one of them.

DR. JALEEL: Since the theme is public reporting and VON's data is not available to everybody, it might make sense to harmonize 478 and 1731, which is already done, and pick that.

CO-CHAIR RILEY: So pick the Joint Commission's?

DR. DENK: I also wanted to --

DR. JALEEL: Is there any objection to that? Are there any downsides to that? I don't know.

DR. DENK: Yes, I also would vote

to decertify or whatever it is the VON, because they don't speak for public disclosure. So why should they be the custodian of the measure?

And I don't care whether it's AHRQ or the Joint Commission, but I don't see why we can't get rid of one of them and just have one measure.

And if VON wants to go and enter into negotiations with the other two parties who figured out how to do it, right, how to harmonize, and if they want to enter into negotiations and harmonize all of them I think that would be great. But I think the burden should be on VON.

MR. CARPENTER: So if I could just say a word. This is Joe Carpenter at VON.

The concerns about extracting ICD-9 data at this point are that we want to be able to differentiate infections reliably. So we have identified specific pathogens that constitute infections as well as a clear

definition of coag negative staph that we consider reliable data.

And the issue that was brought before about the reliability of the ICD-9 data are important to us. We're certainly considering using administrative data in the future and ways that we might do that reliably.

But it may be some time. So I would just say that you're right, we are pointing it toward quality improvement.

And we don't, as an organization, publicly report data but we do encourage transparency. We leave it up to the hospitals. So I just want to make that point. Thank you.

CO-CHAIR RILEY: Thank you.

DR. GREGORY: Yes, I want to make that point too. I think really very strongly as a developer, or someone who believes in quality and someone who does performance improvement, that we need both of them.

So if we're going to reduce it, we should reduce it to two. And I think the reason, you know, being able to do it from both administrative data and from chart audit, it's the same as sort of the randomized control trial, where there's the effectiveness studies and then the real life efficacy studies.

And the data registries will be the effectiveness. And the administrative data are your efficacy. And I think having the opportunity to do it both ways is important.

DR. PROFIT: Yes, I don't necessarily object to that myself.

CO-CHAIR SAKALA: Could we vote, because we'll be here all day?

DR. PROFIT: But could we get a sense of what is it so far? So for AHRQ, we heard yesterday that essentially no extra work is required for the hospital to collect it, right? So I'm trying to still think, there's

no hospital representative, maybe Teri's a hospital representative.

But so for Joint Commission, could you tell us, like what does the hospital have to pay, essentially, for each data point? Or anything, maybe they have to pay nothing and then I think we're all very happy and we can just endorse it.

MS. WATT: The Joint Commission measures can be used by -- go to the website, all of our specifications are publicly available free of charge.

If hospitals choose to collect the Perinatal Care Measure set, and remember this is just one measure in a set of five, and they want that to be reported to the Joint Commission then they contract with a performance measurement system vendor who actually, the vendor provides the data collection tool, which we validate, and all of those good things that you read in this stuff.

And then the hospital collects the

data using that vendor's tool. The vendor, we have a zillion different quality checks that these data are required to be cleaned through, and then the data come from the vendor to the Joint Commission.

There is a charge to the hospital from the vendor, not from the Joint Commission. And that's up to the vendor.

MS. KIEHN: The charge would be the initial setting up of the query. Pulling all of the information, then after that it's really a push of the button. And it goes once we've validated it. So it's not a manual extraction at all.

And I am comfortable with the fact -- because we're going to collect VON no matter what. We'll collect NHSN no matter what. So moving forward with that, I'm fine if we do not recommend this.

CO-CHAIR RILEY: Nancy.

DR. LOWE: Yes, I was just, to move us along, willing to make a motion. That

we keep harmonized 748 and 731, so that's one measure. And that we keep 304, was that the one. Or was it 303?

DR. JALEEL: Take off both, I would say.

DR. LOWE: Take off both?

DR. JALEEL: Essentially.

DR. LOWE: 304?

CO-CHAIR RILEY: Two is a compromise, I think.

DR. LOWE: Two is a compromise. So our harmonized one and 304. So that's a motion.

DR. SIMPSON: But we really have to get rid of one of those two, because there will always be two unless we decide which one we're going to chose and which one we're not, right?

CO-CHAIR RILEY: You said 478 or 1731, we have to pick one. And then 304 or 303, got to pick one of those.

DR. WINKLER: I mean, it's your

decision that you have to pick one. Okay? That's the whole point is, you have to decide.

But certainly, we've heard all the arguments why four is not highly desirable. And there are real issues around coming down to one. As long as you can explain and justify you can pick one, two, three or four.

CO-CHAIR RILEY: It sounds like we have a justification for why you might want to have something other than VON because it's not publicly reported. So that's a reasonable differentiation there.

But then beyond that, we've got to come up and decide -- okay. That's fine. But then beyond that we need to decide whether we feel comfortable having three, which would be 478, 1731 and whichever one, 303 or 304. Or whether we really just want two. And you had a comment?

DR. MURI: Just to state, maybe, the obvious. If 1731, the determination were made to not endorse that measure. Then if

that measure went away from the Joint Commission's stable of measures, that eliminates a whole data pathway.

You know, this is an established method for hospitals to collect these data and to get them publicly reported. And if that goes away, I don't know that there is a similar infrastructure developed for collection of the 478 measure.

DR. BERNS: Or you'd have to start all over. Meaning you'd have to, I mean you could potentially then consider the AHRQ measure. And I'm not quite sure how we got here, but if we chose one and for some reason we chose the AHRQ one --

DR. MURI: Yes. Can I give you a little bit of history here? When we, the Joint Commission, decided to develop or put forward this perinatal care measure set, we chose from the already endorsed, and it was just right after the endorsement, already endorsed perinatal care measures.

Our measure is the AHRQ measure. But here's the thing, and I don't want to say anything to denigrate AHRQ or the AHRQ measures or anything else. We maintain this measure. We update it every six months. It's published, those kinds of things. And that would go away because we would no longer be the measure steward. AHRQ would do whatever it is that AHRQ does, and they do it very well, but the Joint Commission would not be maintaining a measure that we are not the steward for. That's the bottom line, to be blunt.

DR. BERNS: Before we -- I just want to -- we went a little bit past a very bold suggestion that came from two people, Jaleel and Chuck, basically. It would be a bold step for this committee to recommend one measure.

And I feel like we're --

DR. JALEEL: I'm still not clear on why we need two. If you can explain it,

somebody who is opposed to having just one, if they can explain it to me why we need two?

DR. BERNS: You're saying, I think, because from the VON perspective that they're going to collect this, I heard, I think, Dr. Profit say, they're going to do it anyway. You're a member of VON you're going to collect.

It's a great measure and it's 900 hospitals and --

DR. JALEEL: It's a great measure, we have it. And all the hospitals who are the members of VON will get that information anyway. So why do we need a second one for the whole community?

MS. BRANDENBURG: One thing I'd like to remember too. When we originally did this, you know, the difference between the AHRQ and the Joint Commission. The AHRQ one we all passed unanimously right out of the gate, really not much question about it.

The Joint Commission one we

debated for quite a long time. And maybe it's just the way it's written, because we've determined they are very well the same thing. But the AHRQ one seemed to be pretty well cut and dried, easy decision. Where the Joint Commission one, we debated quite awhile about.

DR. ROMANO: This is Patrick. Again, I would emphasize that the difference here is really a difference of two data streams and two co-stewards that are working together.

You know, AHRQ and the Joint Commission have been working together to harmonize this measure. Under NQF policy, we were asked to submit them as two separate measures, even though they've been harmonized, because there are some differences related to the data flow that feeds the measures.

But it would be difficult, I think, to say well, drop one because then, as Ann has pointed out, then you lose a way of accessing that information. Whichever one you

drop, you lose the other method of accessing that information.

DR. BERNS: So I'm confused. Why do you lose that route? It could still be reported. Let's just say the Joint Commission, I'm just -- unless I am misunderstanding it.

Let's say the Joint Commission measure is the one that goes forward, can't the AHRQ measure still be used by whoever wants to use it? What am I missing here? I mean, the measure's out there.

DR. WINKLER: There is a benefit to having it endorsed by NQF. And that's essentially the decision you're making. It's because people, users, will very commonly, something we strongly encourage, that they will select the measures they use in their programs among the NQF-endorsed measures, because it's gone through all of this angst.

DR. JALEEL: Is it possible to have a joint stewardship?

DR. ROMANO: Yes, that forces us to harmonize our measures, which we probably wouldn't do absent this process.

DR. PROFIT: So I find it impossible to chose between the two if the economic effect of both of them is the same as having either one of them, I find it impossible to choose. Like if there's no disadvantage.

DR. WINKLER: I don't think you have to. As long as you have the reasoning. As long as it's not arbitrary or I like red better than blue. I mean, I need something concrete. And you've given it. If you're comfortable --

DR. PROFIT: I mean, for me primarily, you know, if there's no additional work, economic expense for the hospitals. Like who the data flows to is pretty inconsequential to me personally.

DR. GREGORY: Can I just clarify that? It might actually be cheaper, right?

Because you don't need the vendor to use AHRQ, right? You could calculate and report it and not have to go through a third party?

DR. DENK: AHRQ supplies software. This is part of the patient safety indicator thing. And they provide SAS programs, and I don't know what other platforms, for anybody to do this. So yes, and it's free.

CO-CHAIR SAKALA: Can we vote and --

DR. MURI: What you don't have though is the data cleaning that you have with the Joint Commission processes, that's the difference. You have self-reported data being reported. You don't have anybody looking at the quality of those data.

DR. GEE: Reva, can you outline for us, because at this point I think some of us are confused about what we're discussing. Could we just decide what we need to discuss and decide and --

DR. WINKLER: Part of your

discussion is making that decision.

DR. GEE: So my understanding is that we have these two measures, they're very similar. We really can't substantively determine, they've already said they'll harmonize. So I guess who decides, is it AHRQ or, I mean, who decides?

DR. WINKLER: You do.

DR. GEE: Okay. So we have to decide that. So maybe we should vote on that. Or do we have to decide today?

DR. WINKLER: I would appreciate it, because, I mean, is there anything that's going to change if we wait until tomorrow?

CO-CHAIR RILEY: So let's do this. So here's the question. How many people want to keep all four, as they are, on that board and give people the choice?

MS. LESLIE: Can we just do two at a time? There's like two different issues. Can we do the first two and then the second two?

CO-CHAIR RILEY: As long as we can articulate it at the end of this, yes that's great. Okay. Should we have 478 and 1731? Should we have both 478 and 1731, recognizing that they both measure the same thing?

Just raise your hand. Oh, we have this thing. Okay, this is fun, I love this thing.

DR. DRYE: This is Elizabeth Drye on the phone. Can I vote?

CO-CHAIR RILEY: Awesome. You can vote. Hang on one quick second. Jaleel?

DR. JALEEL: If we say yes to this one, then we are going to discuss whether to keep 478 or 1731, is that how it is going to be next?

CO-CHAIR RILEY: No, the question on the table is: do we want to keep 478 and 1731? Do we want both measures on the table? That's the question. Did you get that, Elizabeth?

DR. DRYE: Yes.

CO-CHAIR RILEY: Okay.

DR. PROFIT: So it seemed like the Joint Commission was implying that the AHRQ measure would be lower quality over time? I mean, that's what I heard on the last statement.

You know, that it would be self-reported and not cleaned. And I think that's an important piece, I mean, I think that's not an inconsequential statement, personally, to me.

DR. DENK: I'm sorry. When you said cleaned, you mean that the hospitals' reports will be somehow reprocessed and then they will be published, something other than what they reported?

Or by cleaned you really just mean the kind of analysis you presented here?

DR. MURI: What I mean is that there are 32 quality checks that our vendors put every piece of data that is submitted to them through. For missing data. For, does

this seem like this is aberrant in terms of the volumes that we've seen before and the patient populations?

There's just a number of these things. As well as review for inter-rater reliability of the data and that kind of a thing. That is work that is done by the vendors and then the data are reported to the Joint Commission.

If the vendor finds problems with the hospital's data, they will send it back and ask them to clarify and correct.

There's nothing that goes from the hospital to the Joint Commission that the hospital doesn't know about. I mean, there's not some magic box there.

DR. BERNS: Just a quick question. Patrick, can you remind me how many report this measure to AHRQ? I'm sure it's in this submission, I just --

DR. ROMANO: Well, in general it happens through data sets that are being

collected by state health data organizations or through insurance companies or others, in some cases regional collaboratives that collect this data.

So overall, there are 44 states that collect data of this type from all non-federal hospitals within their states. So there are six states that would not currently have a reporting structure for the AHRQ indicator.

DR. DRYE: What you just mean is that they collect claims data and run the AHRQ stats back on them. The hospitals don't, they just submit their data to the state and this is part of what the states do with it?

DR. ROMANO: Right, I mean technically we don't call it claims data because the state health data organizations are not collecting it for the purpose of processing claims.

But they are generated by the hospitals with the ICD-9 CM codes and they're

submitted to the state health data organizations.

Now, different state health data organizations have different mechanisms for cleaning and validating the data.

So this is where our data stream is a bit different from the Joint Commission's, because the Joint Commission has certain requirements on its vendors that it requires all of its vendors to process the data in a certain way and to validate it in a certain way.

Unfortunately, AHRQ doesn't have the mechanism for doing that. So we are perhaps more dependent on the different data streams that follow in different states. But, that reflects the real world.

As somebody said earlier, there's diversity of different data streams out there in the world.

If a hospital chooses to follow the Joint Commission data stream and to

contract with one of their vendors then that does provide some assurance about the data quality.

DR. DRYE: And if a hospital did that, can they generate the results and give it to their state if the state requires it? I mean, what we're trying to avoid is a hospital having to work to produce the results for two measures.

I'm just trying to figure out. Is that our job to say, well, geez, the State of California shouldn't be requiring this because the Joint Commission is requiring it? It's the same measure.

So I just don't know whether our vote to approve one, you know, if we approve one and not the other we're cutting off opportunities for some kinds of reporting. We just are.

But why should we be -- the measures are harmonized, so if you generate it for JC or Joint Commission aren't you

generating it for the state at the same time?

That was the goal of harmonization, right, you wouldn't have to generate the results twice?

DR. ROMANO: Right. But hospitals are not reporting the AHRQ measure in and of itself. It's a product of applying the software to data that the hospitals are already reporting to state health data agencies.

So that's why it's a different data flow. There's no additional work associated with a hospital using the AHRQ measure.

But there's also, you lose the benefit of the Joint Commission's program for ensuring, or for assessing and evaluating and assuring the quality of the data as it comes into the Joint Commission, because the Joint Commission is directly collecting all of that data through its vendors.

DR. DRYE: Right, and if you

eliminate the AHRQ measure you would eliminate the burdenless way of reporting this measure and you would just have the Joint Commission approach which requires contracting with a vendor more or less, right?

DR. GEE: Could we move to vote on keeping the Joint Commission measure as the single measure? Could we take a vote on that? Are we ready to?

CO-CHAIR RILEY: The vote that's on the table is that we're going to keep both 478 and 1731, recognizing that there are advantages and disadvantages to that.

So everyone can go ahead and vote. So you're voting yes if in fact you believe both should be available.

DR. WINKLER: Elizabeth, you want to vote?

DR. DRYE: Yes, I vote yes.

CO-CHAIR RILEY: How many people should that be? It is what it is, so we keep both. Ten and eight, okay. So then we have

to vote 303 and 304, is that right? Yes.

So the result of that last vote to keep both 478 and 1731 was 10 yes, 8 noes. So now we're moving on to decide, if you vote yes here you're saying you want 303 and 304, so both VON measures in addition to the other two.

DR. ROMANO: Can I ask a question? My understanding from previous NQF processes is that stratified versions of a measure are usually considered part of the same measure.

So we have a number of AHRQ measures that are stratified, where those are sub-measures within a single measures. Would this fall into that classification or not?

DR. WINKLER: Actually, Patrick, this is Reva, you know what, I'm not seeing that in this particular measure. And that's not the way you kind of presented it.

DR. WINKLER: Well, I didn't --

DR. ROMANO: Oh, you're talking

about the VON measure. Well, that was one of the questions we asked, but they are different because the risk models are different.

DR. WINKLER: Okay. Well, I will say, we do have AHRQ measures where there are stratum-specific risk models that are still labeled under the same measure number. So it's just a issue for consistency. Not directly relevant to this panel.

MR. CARPENTER: If I may just say, Joe Carpenter at VON again. Yes, we originally presented two measures because the populations are quite different.

You know, you see an average infection rate in the very low birth weight, it's around 15 percent versus three percent.

But if it's more convenient to include these as separate populations within a single measure, I mean, that could certainly be done.

(Simultaneous speaking.)

MR. CARPENTER: Essentially what

that would mean is that you would still have two separate models, they would just be addressed separately within the measure.

DR. WINKLER: Right. I mean that's something that we could do. It would reduce the measure clutter and the numbers. And if that's something that appeals to you all we can certainly work with VON to do that. Yes, the CLABSI's there too.

CO-CHAIR RILEY: So now we're voting whether we want to keep 303 and 304.

DR. WINKLER: Elizabeth, did you want to vote?

DR. DRYE: Yes, I'm going to vote yes.

CO-CHAIR RILEY: Okay. The vote is 13 no, 5 yes.

DR. WINKLER: You haven't really changed anything. It's more paper. So the reality hasn't changed. We could probably put all the information under one number. The question is it's still two measures,

operationally. So that's the question.

MS. LESLIE: Also you could also vote no on this, meaning, no, you don't want either measure. Or no, you want one measure.

CO-CHAIR RILEY: So the majority voted that they do not want both 303 and 304. So now there's a measure of do we want to keep a VON measure.

DR. WATSON: Did VON, didn't they offer to put them on --

DR. WINKLER: We can. That hasn't changed the measures by putting them under one number. That's something I can do. That's not a really huge deal.

The question is: do you want two measures, because they do have separate risk models. You're talking about two measure results.

CO-CHAIR RILEY: So I think we need to ask the question. Do we want a VON measure at all, is what I'm hearing. A combined VON measure or a VON measure at all,

right? Do we want to combine 303 and 304? That's the question. So if you vote yes then you want the combination, 303 and 304.

DR. WINKLER: Elizabeth, are you still here?

(No response.)

DR. WINKLER: 303 is the measure for all the babies. 304 is the very low birth weight babies.

CO-CHAIR RILEY: But they told us that, and you guys told us that the very low birth weight baby is the one that is at higher risk and the prevalence is higher, that's 304.

Yes, 303, which is the whole population. It's just a yes/no at this point.

DR. WINKLER: So it's 3 yes and 14 no. I think you should take the vote on 304, just to be sure that it's clear.

CO-CHAIR RILEY: Okay. So now we're voting on 304, which is the smallest babies, highest prevalence. Can everybody press their button one more time, because

we're missing one?

DR: WINKLER: I think, you know, we could keep repeating and doing this. The next steps, I think, are important. You guys have made decisions and recommendations on all of the measures that we can capture and, in fact, are in the process of capturing.

And essentially our next step is to bundle it all up and post it for public comment. You're going to get a chance to see what that feedback is. And at that point you may need to rethink some of the issues based on the feedback you get.

It'll be an opportunity to take one more look at it. And the fact that some of these are close, yes, these are tough. These are not easy issues and that's kind of the reality of the world we're working in.

So what we'd like to do, as everybody is leaving, just a couple of followup things. We'll have a summary out to you shortly, probably early next week for you

to take a look at.

One area we didn't get to on the agenda, although it came up a couple of times. If your thoughts about the kinds of measures that aren't currently in the portfolio, the gaps, the kinds of things you'd like to see developed, shoot me an email.

Send this in, we'll include them, because it's a nice paragraph to provide guidance to the world about the kinds of things that, you know, we looked at these and it's okay but gee, we wish we had measures of this, this and this as well. So that's an important thing.

You can send it to everybody, yes. I mean, it won't go anywhere in NQF if you don't include me. So you can talk among yourselves all you want to, but if you want us to put in the report it's got to come to me.

So I think that's really all the business we had. Thank you all for -- we need to do public comment --

MR. CARPENTER: Can I just ask what the vote was on 304?

DR. WINKLER: It was yes 9. No 8.

MR. CARPENTER: Okay. Thank you.

CO-CHAIR RILEY: Sure. So is there anybody on the phone, nobody's still in the room, who wanted to say anything? There's somebody there behind the pillar, I'm sorry.

MS. JOHNSTON: I'm Tina Johnston, I'm from the American College of Nurse Midwives. And I want to thank you all for all of your hard work.

And I just wanted to call to your attention a landmark, unprecedented document that was just released today, that just came across our emails, that involves the American Academy of Family Physicians, AAP Pediatrics, ACNN, ACOG, A1, SMFM and the osteopath OB/GYNs.

And that document is entitled, "Leading Healthcare Organizations Issue Recommendations for Quality Patient Care in

Labor and Delivery. An unprecedented collaboration creates a joint call to action."

And basically, I will forward this to anybody that wants it, but it's going to be on all of our organizational websites.

We recommend, for healthcare providers and administrators, that we ensure patient-centered care and safety. That our organizational priorities, guiding decisions for policies, fostering a culture of openness by promoting active communication of good outcomes and opportunities for improvement.

Developing forums to facilitate communication and track issues of concern. Providing resources for clinicians.

And this is the one that I think is key and most relevant here. Providing resources for clinicians to be trained in the principles of teamwork, safety and shared decision-making.

Develop methods to systematically track and evaluate care processes and

outcomes. Facilitate cross-departmental sharing of resources and expertise. Ensure that quality obstetric care is a priority that guides individual and team decisions.

Identify and communicate safety concerns and work together to mitigate safety risks. To disseminate and use the best available evidence, including individual and hospital-level data to guide practice patterns.

And so as we're all discussing this today, this is coming out. I think that one of the major gaps we have here is a way to really track the processes that are going on.

And I think that maybe moving forward, if you think about this in terms of systems and in terms of teamwork, communication, and those processes that we know produce good outcomes.

Moving forward, I think that there are major quality gaps in that arena that these documents address. So thank you for

listening.

DR. WINKLER: I look forward to seeing the document. Anybody else? Anybody on the phone? Casey, is there anybody? Are the lines open? Okay.

OPERATOR: I can open the lines now.

DR. WINKLER: Okay. Please do.

OPERATOR: The lines are open.

DR. WINKLER: Any comments or questions from anybody on the phone? Another comment?

PARTICIPANT: Yes, I might have missed it. Did you talk about research recommendations for the future for measure developers?

DR. WINKLER: We just did. There wasn't time for the conversation but during the course of conversation, things have picked up and I've asked everyone to send me their thoughts and ideas.

PARTICIPANT: Okay, good. Can you

say when you need that by?

DR. WINKLER: Within the next week.

(Whereupon, the above-entitled matter was concluded at 3:57 p.m.)

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

Before: NQF

Date: 11-30-11

Place: Washington, DC

was duly recorded and accurately transcribed under my direction; further, that said transcript is a true and accurate record of the proceedings.

Neal R. Gross

Court Reporter