NATIONAL QUALITY FORUM

+ + + + +

PATIENT SAFETY 2015-2017
STEERING COMMITTEE

+ + + + +

THURSDAY
JULY 28, 2016

+ + + + +

The Steering Committee met at the National Quality Forum, 9th Floor Conference Room, 1030 15th Street, N.W., Washington, D.C., at 9:00 a.m., Ed Septimus and Iona Thraen, Co-Chairs, presiding.

PRESENT:
ED SEPTIMUS, MD, Texas A&M University Health Science Center; Hospital Corporation of America; Co-Chair
IONA THRAEN, PhD, ACSW, Utah Department of Health; Co-Chair
JASON ADELMAN, MD, MS, Montefiore Medical Center
CHARLOTTE ALEXANDER, MD, Memorial Hermann Medical System
KIMBERLY APPLEGATE, MD, MS, FACR, Emory University
LAURA ARDIZZONE, BSN, MS, DNP, CRNA, Memorial Sloan Kettering Cancer Center
CHRISTOPHER COOK, PharmD, PhD, bioMerieux
MELISSA DANFORTH, The Leapfrog Group
MARTHA DEED, PhD, Patient Safety Advocate
THERESA EDELSTEIN, MPH, LNHA, New Jersey Hospital Association
LILLEE GELINAS, MSN, RN, FAAN, CHRISTUS Health
STEPHEN LAWLESS, MD, MBA, FAAP, FCCM, Nemours
LISA McGIFFERT, Consumers Union

SUSAN MOFFATT-BRUCE, MD, PhD, The Ohio State University
PATRICIA QUIGLEY, PhD, MPH, ARNP, CRRN, FAAN, FAANP, Nurse Consultant
MICHELLE SCHREIBER, MD, Henry Ford Health System*
LESLIE SCHULTZ, PhD, RN, NEA-BC, CPHQ, Premier, Inc.
TRACY WANG, MPH, Anthem
KENDALL WEBB, MD, FACEP, University of Florida Health Systems
ALBERT WU, MD, MPH, FACP, Johns Hopkins University
YANLING YU, PhD, Patient Safety Advocate

NQF STAFF:
HELEN BURSTIN, MD, Chief Scientific Officer
ANDREW ANDERSON, MHA, Senior Project Manager
JASON GOLDWATER, MPA, Senior Director
ANDREW LYZENGA, MPP, Senior Director
ELISA MUNTHALI, MPH, Vice President, Quality Measurement
JESSE PINES, MD, Senior Director
DESMIRRA QUINNONEZ, Project Analyst

ALSO PRESENT:

JAMIE FOX, APNP, Children's Hospital of Wisconsin
THERESA MIKHAILOV, MD, PhD, Children's Hospital of Wisconsin
PAM OWENS, PhD, Agency for Healthcare Research & Quality (AHRQ)
DANIEL POLLOCK, MD, Centers for Disease Control & Prevention (CDC)
TOM RICE, MD, Children's Hospital of Wisconsin
PATRICK ROMANO, MD, MPH, University of California, Davis
RAMESH SACHDEVA, MD, PhD, JD, AHRQ-CMS CHIPRA Pediatric Measurement Center of Excellence (PMCoE)
DONNA WOODS, MD, Pediatric Consultants, LLC

* present by teleconference

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

Welcome, Recap of Day 1 5

Consideration of Candidate Measures (Cont.)

3005: Initial Risk Assessment for Immobility-Related Pressure Ulcer within 24 Hours of PICU Admission.....16

3006: Initial Baseline Screen of Nutritional Status for Every Patient within 24 Hours of PICU Admission.....51

2983: Potassium Sample Hemolysis in the Emergency Department.....118

NQF Member and Public Comment 172

Consideration of Candidate Measures (Cont.)

3025: Ambulatory Breast Procedure Surgical Site Infection (SSI) Outcome Measure.....174

0450: Perioperative Pulmonary Embolism or Deep Vein Thrombosis Rate (PSI 12).....207

2909: Perioperative Hemorrhage or Hematoma Rate (PSI 09).....250

Next Steps/Committee Timeline 264

NQF Member and Public Comment 268

Discuss Gaps in Measurement 268

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

9:04 a.m.

CO-CHAIR SEPTIMUS: Okay, so, welcome back, everybody. We had a full day yesterday. We got through all the measures on the agenda. How about that? So we only went over 30 minutes which I think was pretty good.

And so we really had, just to kind of recap things, and Iona can fill in what I don't say, we had breakfast. Come on guys, lighten up. It's day two.

We went through some new measures. In fact, most of the measures we considered yesterday as you remember were new. A number of them were process measures. Some were outcome measures.

We looked at potentially harmful drug interactions, med reconciliation and dialysis, which we felt was a really important issue. And the developers were kind enough to follow up with someone that knew a little bit more about the statistics and the proposal so we could actually act on it. And that was extremely useful.

In the afternoon we had some PACE measures that we went through. The first one on pressure ulcer prevalence we declined to endorse for good reasons.

I did talk to the developers after that. I think they got some feedback as to how to adjust that.

We did pass the next two on fall rates and fall rates with injury. But I think the developers learned a lot about the process and what they needed.

We had two of our outstanding nurses who had to recuse themselves and they're here this morning. Not to talk about the measure.

But I would like, and I don't want to put anyone on the spot. I really think that PACE has phenomenal value to healthcare delivery.

So if one of you would just like to talk, not about the measure. You can't talk about the measure. But just talk about your viewpoint of the PACE program so when some of these things come back we have a better understanding about what it's

actually doing.

Because I think we got the sense that it's really doing a lot, but the measure, especially the first measure, didn't match up to what they needed to do.

MEMBER EDELSTEIN: Good morning, thank you. I am from the New Jersey Hospital Association as you all know.

And in New Jersey we represent the PACE organizations in our state. So I'm very familiar with the model.

It is, as you heard yesterday it is a fully capitated model of care that serves the frail elderly 55 and older who qualify for nursing home placement.

Most of the participants live in the community. They participate in the PACE center usually two to three days a week.

In addition many, not all, but many receive home care services through the PACE organization. So PACE is both provider and insurer which makes it unique in the provider

community for sure.

Most of the participants as I said come to the center two to three days a week which means that the interdisciplinary team, at least some members of it if not all, lay eyes and sometimes hands on those participants more than once a week.

And they see them in their home, they see them in the center. The drivers of the transportation vehicles who bring the participants to and from PACE are a vital part of identifying changes, of understanding what's going on in that person's life beyond their healthcare.

And it really stands out as CMS's first really fully integrated dual eligible arrangement for this population.

It is being held out actually as a model for all of the fully integrated dual eligible models that CMS has put forward in recent years.

So, beyond the IDT team itself which is so critical to planning appropriately for the care in every aspect of a PACE participant's life there is a lot of care delivery going on every day.

PACE participants have to -- I'll use the word relinquish. It's a little strong word, but they have to relinquish their relationship with any prior physicians and have the PACE physicians as their primary care providers.

They get all of their primary care through the clinics in the PACE center. Specialty care, same way.

And whenever they need home care, or hospital care, or nursing home care, the PACE organization must contract for those services.

They can provide home care directly and most do. But they can also contract for home health.

As was mentioned yesterday most of the home care provided to PACE participants is personal care assistance in their own homes.

But when a PACE participant does need a higher level of care temporarily the PACE organization pays for it, contracts for it, oversees the quality of it and remains integrally involved in the care planning process throughout

an acute care stay or a long-term care stay.

So hopefully that gives you a little bit more of a feel for sort of the all-encompassing nature of PACE.

And beyond care delivery there's also all of the intangibles like pest control, and air conditioning units, and deep cleaning of an apartment, and removing bed bug infestations, and all of those things that are social determinants of what can happen to a frail older person living in the community if they're not attended to.

CO-CHAIR SEPTIMUS: Thank you for that explanation. Lillee, I don't know if you wanted to?

MEMBER GELINAS: Thank you, Theresa, that was terrific. And thank you, Ed. Helen, welcome. We're sorry you missed this.

But first of all I was emailing Andrew and Jesse yesterday because I knew I'd have to recuse myself, and couldn't vote, and couldn't even talk in open comment.

I have a great respect for NQF for

holding our feet to the fire for that. As a matter of fact, if not I think the opportunity that we have before us would be suspect.

So I wanted to make sure all of you were quite aware. I was aware of that before I arrived.

From the standpoint of the measures we were considering I want to commend this committee. I thought you were incredibly professional to go through almost three hours of what I thought was a very disappointing presentation.

Last year when we had to present the nursing measures on falls and pressure ulcers, the ANA measures, Pat and Victoria Rich and I actually came into Washington ahead of time, two days ahead of time. We prepped. We were ready.

We did everything we could to make sure we were anticipating your questions.

And we were just sitting here yesterday saying my, my, for those of you that may not understand the world of nursing measurement this was a real disappointment.

And so I want to thank you for your

professionalism. I think you did the right thing by sending the pressure ulcer measure back to the measure developer.

But I am hopeful that it will come back and will close the gap. Because at the end of the day no measures are perfect, and at the end of the day if we're not measuring then we can't hold providers and consumers accountable for care. So, I'm hopeful that the homework will be done.

But from a social standpoint I have to tell you what an honor it is to serve on this committee. The chemistry on this committee is absolutely amazing.

We all serve on a whole lot of committees, every one of us, and this one I see an awful lot of heart and soul, and a lot of work behind the scenes.

So, I just want to publicly thank you for your professionalism yesterday. I thought you were spectacular and I'm very proud to call you committee member. Thank you, Ed.

CO-CHAIR SEPTIMUS: Thank you. Lisa?

MEMBER MCGIFFERT: Before I worked on patient safety I worked a lot on disability issues. And this program was really I think emerged from the disability community. And yet it's where the elder community, elder activists came together.

It really does represent the kind of thing that I think a lot of us would like to see in the future as an alternative to nursing home care.

And when you think about all the money that we put into nursing home care, and then you think of what it takes to keep somebody at home it does involve all those different things, healthcare as well as social services.

And this program and others like it I think are really critical for the future of our system.

And so I was glad to see that it was being brought forward, and hope to see some more measures for that later.

CO-CHAIR SEPTIMUS: Fantastic. Okay. Appreciate those comments.

So in the afternoon we went through what I think everybody agrees was a public health, I don't want to say crisis, but public health concern about opiate overuse. And we approved all three of those measures.

So that was I think very important. These were all process measures and these were all first-time. So we hope that when they come back to us in three years that they will have data to show its effectiveness in terms of monitoring high use of opiates in the community.

Then we had public comment. We had an excellent comment from someone who was actually here at the end of the meeting.

So for today since I know some of you have to leave by 2 we're going to try to get through the measures that we can.

We've been trying to reach out to the developers to see if we could move up some of the discussion before discussing the gaps in measurement.

And so we may be adjusting the schedule

so that we can accommodate everyone. We don't want people to leave and then not have a quorum.

We'd certainly like to be able to finish all our work. This would be the first for the patient safety committee to finish all our work before we adjourn at 3 o'clock today without having to get a follow-up call. So I think that's our goal.

I think the work that you all do, and I'll second what people have said. This is an incredible committee with incredible experience.

One of the things I said yesterday and it came out again last night at dinner is I think we like each other, and I think we've really bonded as a committee. And I think that's a real credit to all of you.

I can't say enough about Iona as co-chair who keeps me in line.

And of course the heavy lifting behind the scenes is done by the NQF staff. We could not do our work without them.

(Applause)

CO-CHAIR SEPTIMUS: Andrew's still awake. He said his child slept well last night. Okay, so is any of the developers here?

Okay, so if it's okay with you to kind of adjust the agenda let's go to 3005 and 3006, initial risk assessment for immobility-related pressure ulcers within 24 hours of PICU admission. And then the next one, we'll go to the next one. So why don't we start with 3005.

And I think, Pat, you were going to lead that discussion after the developers, correct? Is that right, Pat?

CO-CHAIR THRAEN: It's the initial risk assessment for immobility-related --

CO-CHAIR SEPTIMUS: 3005. So, if the developer could come up that would be great.

CO-CHAIR THRAEN: And Steve, you're back up and Martha.

CO-CHAIR SEPTIMUS: And so, just to let everybody know this is an eMeasure. I think last year was the first time we had an eMeasure.

(Simultaneous speaking)

CO-CHAIR SEPTIMUS: It's a process eMeasure and it is a new measure. So are the developers here? Come up front. We're a friendly group.

CO-CHAIR THRAEN: So while they're working on that Ed didn't give me a chance to talk.

CO-CHAIR SEPTIMUS: No, I was going to say. You can see how well we work together.

CO-CHAIR THRAEN: I just want to make a comment that I've had several people come up and comment to me about the level of comfort and camaraderie and respect that people are feeling in this process.

And I just wanted to feed that back to you, that I think that we have gotten to a place where we honor each other's expertise. It doesn't mean we always agree, but we can honor it and also articulate the disagreements and support each other in that process.

So, we feel like we have a good complementary group of people. It's not competitive and at the same time we can be

supportive of each other while disagreeing and listening to each other's point of view.

So I just want to honor you all for that. That's what you bring to the table as well. Thank you. Now I've said my piece.

MEMBER WU: The chairs are always responsible for that, setting the tone.

CO-CHAIR THRAEN: Thanks, Wu. Thanks, Albert. Dr. Wu.

MR. LYZENGA: So I think the developer is still trying to kind of assemble their team, so maybe we could have Jason Goldwater who works on our eMeasure team here come up and say a few words.

CO-CHAIR SEPTIMUS: Hi, Jason. Jason gave us a great discussion last time introducing us to eMeasures so we appreciate you coming back.

(Simultaneous speaking)

MR. LYZENGA: And there's also one measure that we're going to be considering that is eligible for trial use approval. So Jason will say a little bit about that as well.

MR. GOLDWATER: Right. So good

morning, everyone. Always a pleasure to be here even though I'm a little fatigued due to the long-winded nature of the Democratic Party. Oy vey. And there's still one more to go tonight.

I sort of have to. I am somewhat compelled to do so because I have a wife that's going to do so, and it's that or America's Got Talent and that's not on tonight. So.

So, what I want to do is just talk about a couple of things. One is the way eMeasures are brought into NQF and how they are generally examined before they get to you, and things for you to consider.

And then to talk about the trial use program which is going to be something that will be considered today.

EMeasures have certainly evolved significantly over time. I know there are a number of you that can think back to the good old days when CMS used to be called HCFA and we were doing manual chart abstraction for quality measurement which is not to say that still doesn't go on from time to

time.

And there was a real push in the early two thousands to move out of chart abstraction into using electronic health records to populate quality measures automatically and get out of the abstraction in order to reduce the amount of human error that could occur during the abstraction process, and also to ensure that standardized codified data could be used in the measurement itself.

In 2003 CMS ambitiously started a project known as the Doctors Office Quality Improvement Technology bracket or DOQ-IT for short. Some of you may remember that.

I had the -- how can I put this delicately -- the honor of being the project director for that initiative which failed miserably.

And not because the intentions were bad, but because EHR adoption across hospitals and physician offices in 2003 was less than 20 percent.

And then HITECH passed and suddenly

there was a large influx of money. And suddenly now as we enter -- or we're halfway through 2016 hospital EHR adoption is well over 80 percent and physician office EHR adoption is almost at the same level.

And so now we have started to revisit the idea far more aggressively in the utilization of eMeasures as opposed to those that are chart abstracted.

In the good old days back in the beginning of quality measurement most of the eMeasures that would come in were de novo, brand new that were created using specifications that they could find in EHRs.

That is not the case anymore. Because again, when they first started there were not a lot of EHRs so the project was for most intents and purposes not done very successfully.

EMeasures now come into NQF in one of four ways. The first is a de novo measure which everybody knows. So it's a brand new measure that is being created for patient safety using data that

is from an electronic health record, preferably structured data. It's a measure that is not existing at the moment, and it is a measure that we need to look for consideration because preferably it is filling a current gap in patient safety.

The second way it can come in is what we call a re-specified measure which is a measure that is currently a chart abstracted measure.

And the desire by CMS is to move away from that and make it into an electronic measure.

So they take the specifications of the chart abstracted measure, map it to the same data elements found within the EHR, re-specify it and send it to us.

The third way which is Elisa's favorite way of a measure coming into NQF is what we call a legacy measure.

That is a chart abstracted measure that is currently used in a federal program such as PQRS, or the Meaningful Use program, or IQR.

And it is chart abstracted and the

desire is to make it an eMeasure. So it's already being used in a federal program and has probably been done so successfully, but now they don't want it to be chart abstracted, they want it to be e-specified.

And so they do the same thing. They will then map the chart abstracted elements to the elements within the EHR and submit it to NQF.

And then the final way it can come into us is through trial use which I'll explain in a moment.

Any type of eMeasure that comes into NQF has to be tested in at least more than one EHR, or essentially two. It's all in the wording. It never gets new. We always say more than one and people are like oh, two. Yes, two.

So, it needs to be in at least two EHR systems.

And it has to be different systems. Now, some of the questions that we get is, well, I'm testing it at Cleveland Clinic, and then I'm testing it at Memorial Hermann. They both have

Epic. Is that construed as two separate EHR systems?

And the reality is it is. If you look at the way the implementation has been done at Cleveland and the way it has -- and mind you I'm just throwing these hospitals out -- has been done at Memorial the implementations may be different. And so subsequently that would be construed as two different electronic health record systems.

And given that Epic roughly has almost 35 to 40 percent of the marketplace it's not unreasonable to look at two Epic systems as being two very separate and distinguishable electronic health record systems.

It has to be in the appropriate format in that it has to be in the health quality measures format which as measure developers know if you're developing an eMeasure most of them use the measure offering tool which was originally created by NQF which has now been taken over by the MITRE Corporation.

And once that is developed the

appropriate format is created.

It has to map to what we call the quality data model which is in other words high-level elements that specify what each component of the measure is. Race, ethnicity characteristics. This is an encounter. This is a diagnosis. This is a procedure.

And then that way at least you're standardizing how the measure is laid out so that when you're implementing it into your EHR system you know exactly what codes you have to map to where.

The other part is in addition to it being formatted correctly it also has to contain value sets. And those value sets are at this time maintained by the National Library of Medicine and their value set authority center.

A value set is really a building block of a measure. It's a coded element that represents a condition or a diagnosis. And it maps to a nationally recognized terminology like ICD-9 or ICD-10 or CPT or RxNorm, or codes that are used

pretty commonly throughout most systems.

The value sets have to be published. Believe it or not there are developers, and I'm not saying that's you, that create their own value sets and then don't publish them in the value set authority center. Which means nobody can use them other than the developer themselves.

So, we recommend -- well, actually we don't recommend. Now we just demand that you have to be able to publish them so everybody can see them.

Once that's all done then the measure comes to us. And we also have to look at feasibility. Just like you would look at feasibility on any type of measure you also look at feasibility for the eMeasure.

And the things that you look at and to consider when examining an eMeasure are is the data available which means is the data in an EHR system.

And is the data structured. In other words, is it a coded element that can easily be retrieved from an EHR.

So, for example, most physicians or nurses will input information in the EHR and there are structured fields that they input that information into.

But occasionally, and actually I should a little bit more than occasionally, when they have to talk about follow-up plans, or specific instructions for a patient they don't code those elements. They type them in the free text fields.

And if that's part of the measure that becomes pretty difficult to get out because every EHR is different in where that information is actually stored. So, it's something to keep in mind.

The second is is the data using a national standard, or a national vocabulary. Because if it's using something that you don't know or haven't heard about, and that's very, very rare, then the idea that it could be implemented across many EHR systems is incredibly restrictive.

If it uses a national code like a SNOMED code, or an ICD code, or a CPT code that's what

everybody generally uses. That makes implementation a little bit easier.

It's also worth noting that is the use of the eMeasure when it's implemented really interrupting work flow. Do you have to take 20 minutes to input the information into an EHR for this measure because that's 20 minutes you're not spending with your patient?

And there's no way that NQF or anybody would really want an eMeasure that requires so much time away from the patient that it actually becomes more burdensome to do than actually chart abstracting the measure.

The idea of eMeasurement is to make it easier to get the information into a measure rather than through the chart abstraction process.

Now, there's one small caveat. Of course there is. Which is these legacy measures which I just talked to you about. Right. I roll my eyes too when I get those. It's like really? We're doing this again?

But legacy measures are already in

federal programs, they're already being chart abstracted and now they want to make them eMeasures.

A lot of times it's really hard to get data from two EHR systems to test that because these have been chart abstracted measures. They don't have EHRs where these have been implemented unless they've just done this on their own which rarely happens.

So what do we do when that occurs? One of the solutions we've come up with which I will come out and say is not a permanent solution, but it is one that we are currently using, is that developers can simulate a test data set of patients to evaluate the logic of a measure to make sure it calculates correctly, it's producing the right metric, without actually using an EHR system.

There is a tool that MITRE created called Bonnie. I always get asked this question what does Bonnie stand for. I have absolutely no idea. I don't think it stands for anything. In all honesty it's probably the daughter of the

developer or the pet. Developers are very fond of naming things after their children or pets. Speaking as a former developer. It's just a lot easier to do. You don't have to be creative.

So they use Bonnie to create a simulated patient -- synthetic patient test deck of 50 to 60 patients that would meet the criteria of the measure.

And it's important if you see this that if they've created a synthetic patient test deck in Bonnie that it actually represents a population of patients you would actually see.

Like you don't want to see everybody meets the measure. You would want to see people that are excluded. You want to see people that are included. You want to see people with different conditions to make sure the logic of the measure calculates correctly.

And if so, while it is not a complete, absolute pass on feasibility or reliability or validity, it is safe to say to some extent that if the logic is calculated and the metric is accurate

that in implementation there is a strong probability that the measure would actually work.

The last thing --

CO-CHAIR THRAEN: Before you go on, I want to ask a clarification question on Bonnie.

MR. GOLDWATER: Sure.

CO-CHAIR THRAEN: So, when we were looking at binomial Bonnie testing we were seeing really high 97s, 0.9397. Is that addressing the issue of the ones that were -- I want to make sure I understand.

So you were saying that the simulation patient set should include some that don't belong.

MR. GOLDWATER: That's correct.

CO-CHAIR THRAEN: And that your measure shouldn't be 0.999.

MR. GOLDWATER: No, no.

CO-CHAIR THRAEN: So you can discriminate.

MR. GOLDWATER: That's correct. Right.

CO-CHAIR THRAEN: Okay.

MR. GOLDWATER: If they come back with results that are 100 percent there's a problem because that's not accurately resembling -- well, because you have to be able to show that there is at least some ability of the measure to calculate correctly for those that don't meet the criteria for inclusion.

Because if you were to test it in the EHR that's what you would be testing.

CO-CHAIR THRAEN: I think Andrew's got a clarification.

MR. LYZENGA: That's different from the binomial model of doing the signal to noise analysis.

MR. GOLDWATER: Right. All right, so the last is trial use which you're going to be hearing today from my very dear friend Michael Thelon at some point.

And trial use was a program that was brought into existence at the beginning of last year.

And the reason it was developed was

because we still have significant gaps in measurement. I don't think I'm saying anything that all of you don't know.

And there is the need to facilitate the development of measures that can fill those gaps. But testing of these eMeasures at times can prove to be extremely burdensome and very difficult to do, particularly when you're looking for specific data elements, making sure those elements are structured, or finding ways of mapping free text elements into a structured format.

So there was two things we could do. We could completely ignore the development of those measures because they were not going to be able to meet the criteria, or we could alter the criteria for those measures specifically which in a way is a slippery slope because then you've got to start altering the criteria for others.

So, the trial use program was created.

When a measure comes before you that is being considered for trial use the measure is to be evaluated the same way any measure would be. Is

it reliable. Is it valid. Look at the feasibility.

The feasibility will use Bonnie as I've just described, and in some cases they may be able to actually use real data also to test the measure in conjunction with Bonnie. Not everyone, but some do.

If you agree that this measure meets the criteria of NQF the measure does not get endorsed. The measure gets put into the trial use program.

And what happens at that point is the measure is then put into the field and is implemented in some sites and data is collected while it's in the field. So essentially it's being tested while being used.

And after a period of time that the developer has collected enough data, while that measure has been used, they then take the measure out of the program. They evaluate the measure as they would if they had actually conducted testing.

They bring it back before you and you reconsider the measure again with those testing

results to see if it can be considered for possible endorsement.

CO-CHAIR SEPTIMUS: So, our usual process, and NQF staff, correct me, is usually it gets endorsed just generically and then it comes back in three years for re-endorsement.

So, would that be the same with these trial use?

MR. GOLDWATER: No.

CO-CHAIR SEPTIMUS: Or can they come back the next year?

MR. GOLDWATER: They could come back the next year, yes. It doesn't have to be a three-year period, no. It can be -- they have to up to three years to do it. That's correct.

CO-CHAIR SEPTIMUS: So they're not really being endorsed, they're being given --

MR. GOLDWATER: Entry into the program.

CO-CHAIR SEPTIMUS: The opportunity to test it. But then they still have to come back within three years.

MR. GOLDWATER: Yes.

CO-CHAIR SEPTIMUS: And then they can get full NQF endorsement.

MR. GOLDWATER: That's correct.

CO-CHAIR SEPTIMUS: Okay.

DR. BURSTIN: Just part of the logic behind it was we didn't want to hold up innovative measures because the EHR systems weren't quite there yet to be able to test them.

So get them to market. Label them in a way people know they're not completely ready for prime time but please try these, explore these so that they can potentially get ready for prime time.

MR. GOLDWATER: Right.

CO-CHAIR SEPTIMUS: So, if we decide through going through our usual process of evidence gap, et cetera, we don't think it's quite there yet even now then they don't move forward?

MR. GOLDWATER: That's correct.

CO-CHAIR SEPTIMUS: Okay.

MR. GOLDWATER: But what I want to emphasize is it's not being considered for

endorsement. It's being considered for entry into a program so it can be tested into the field.

(Simultaneous speaking)

CO-CHAIR SEPTIMUS: And we'll make that clear with each of these eMeasures that we're going to be discussing this morning.

MR. GOLDWATER: Correct.

CO-CHAIR SEPTIMUS: Steve, you had a question? And then Charlotte.

MEMBER LAWLESS: Yes, I want to compliment you. You actually made it so very, very clear.

MR. GOLDWATER: Oh, thanks.

MEMBER LAWLESS: And so I'm waiting for the next convention to hear you.

It's so clear. Is this conversation you have in a document or something that people get so they actually can say this is the process you have to go through? I mean is it ahead of time?

MR. GOLDWATER: We do. Yes.

MEMBER LAWLESS: I really recommend it actually because I get asked a lot how to develop

a measure. What you described is what on the e side we go through it all the time.

I would really ask NQF either to put it as a webinar, or a course, or a book.

MR. GOLDWATER: So we've done the webinar. I'm happy to do it again. And we do have a document that pretty much describes everything I've just talked about.

And we're actually revising that at the moment. Reva and I are revising that at the moment. And I think it will probably be out in mid-August.

MEMBER LAWLESS: So I think it's applicable actually for internal development.

MR. GOLDWATER: That's correct.

MEMBER LAWLESS: So thank you.

MR. GOLDWATER: Absolutely.

CO-CHAIR SEPTIMUS: Jason, could that information -- I mean, some of us have been involved in NQF for awhile and may or may not continue our roles at some level.

Would it be something that you could

share when it's available --

MR. GOLDWATER: Absolutely. We'll probably let everybody know once it's done, it's released because it will be sort of a list of the scenarios of the way eMeasures come in, how eMeasures are evaluated, the things to consider. A new feasibility scorecard. A variety of things. So yes, we will release all of those.

CO-CHAIR SEPTIMUS: Charlotte and then Lillee.

MEMBER ALEXANDER: So is your trial measure program only for eMeasures?

MR. GOLDWATER: Yes.

MEMBER ALEXANDER: Because we've had some measures come through that didn't have the data, needed to be out there to get the data. I'm thinking of some of the radiology measures that came through a year or so ago.

And it seems like there might be an opportunity for NQF to provide that type of support for other measures as well.

MR. GOLDWATER: Well, I think that's

something to potentially consider. For now it was really about eMeasures and that was also largely driven by the market itself with the desire, as Helen put it, to create a lot of innovative measures that were filling needed significant gaps in care.

And rather than holding the process we came up with the trial use process as a way of at least moving those out to see if they work.

CO-CHAIR SEPTIMUS: Before we go to Lillee is there anybody on the phone that's part of the committee that has not announced themselves? I should have done that earlier and I apologize.

MEMBER SCHREIBER: Hi, Ed. It's Michelle Schreiber. I'm on the phone.

CO-CHAIR SEPTIMUS: Hey, Michelle. You get a gold star. No, I think I've risen that. You're a platinum now. Two days in a row.

MEMBER SCHREIBER: I appreciate being allowed to participate by phone. I couldn't be there in person.

CO-CHAIR SEPTIMUS: Okay, thank you. Lillee, go ahead.

MEMBER GELINAS: Thank you. And I agree, incredibly articulate.

The pain on the provider side around the workplace impact of EHR implementation. And some of the data that we track about the nursing work environment.

Nursing turnover is extremely high in this country right now. And unfortunately when you see exit interview after exit interview of why nurses are leaving they're so constrained at nursing the computer they're not nursing the patient.

And we've done some studies showing that 80 percent of time is spent on documentation burden, not nursing the patient type thing.

Conversations with vendors don't go well. We have MEDITECH, Cerner and Epic and they don't talk to each other. And we have Midas for quality data extraction. And I'll go on and on.

The number of FTEs that we still have for manual chart abstraction is phenomenal.

So, give me your hope trajectory on how

quickly that all of these stars are going to align. Because I'm very hopeful that they will. I think that a lot of the energy and synergy is well underway, and the practical and tactical is now beginning to take hold whereas before it could have been pie in the sky.

So our crystal ball is still pretty foggy, but we're moving fog to concrete now.

But do you have any hope whatsoever that we're going to really move to vendor-to-vendor interoperability at a level I think it was originally conceived by the American health information community, and Secretary Leavitt, and everybody else?

Is there anything that we can do and NQF can do to really push the market to make interoperability real?

Because we can have all the best measures in the world, but if we don't have good interoperability it's still going to be painful to get the data.

MR. GOLDWATER: So, it's great that you

asked a very simple question that will take two minutes to answer. I'm kidding. We could have a conversation about this literally for the rest of the afternoon.

CO-CHAIR SEPTIMUS: I can give you three minutes.

MR. GOLDWATER: Thank you, Ed. I'll answer this as succinctly as I can.

I think that interoperability is the major barrier here. It always has been. It's been a barrier for 25 years.

There's been slow, incremental progress. There's been a lot of discussion about removing the information blocking to allow for better sharing of information and data.

There certainly does seem to be a degree of willingness at least publicly by vendors to do this.

I'm somewhat cynical as some of you are that I've been down this path a lot. I've often joked ONC has created a roadmap for interoperability and this is the fourth roadmap

that they've created in the last 15 years.

So I think that for now we can sort of get to a point where we can use these more effectively because you can look at things such as if you're using structured elements they're obtainable in the EHR. It does not force you to have to do some degree of text mining or in-depth examination about where those fields are.

And when I review eMeasures that's the first thing I look at which is is the data readily available.

Like I know having looked at a gazillion systems over the years that you know the elements that are there. You know the elements that are not.

And if the elements are there, and they're structured, and they're available it's hard to assess the actual impact on workflow because I've been to hospitals that have an Epic system.

They've got terminals every 10 feet and I'm still watching nurses write down on paper

what's going on and then double input into the EHR which is incredibly ineffective.

So I think that there is a way that we can at least help the problem.

Are we going to get to full interoperability? NQF is certainly helping in this regard because we're working with ONC on how to effectively measure interoperability in order to understand what the problems are and how those problems can be solved.

A lot of this is really going to take shape in the next couple of years about what's going to happen. You know, is MACRA really going to be sort of the driver that opens up these systems. Are there really going to be significant penalties for those who continue to block information.

Are we going to get to a national standard across all terminologies and vocabularies that everybody will use?

And most importantly, are we going to get to a way where we can uniquely identify a patient? Because all the interoperability talk in

the world is great, but if we can't specifically attach that to a patient that poses a problem for quality, and safety in particular.

When I hear discussions of, well, we have probabilistic algorithms. We can get a 97 percent chance of getting it right. You know, that sort of gives me a lot of pause because there's 3 percent of the patients are going to be wrong. And this is healthcare. You don't get this wrong.

So, I think you pose an incredibly great, philosophical, in-depth question that I would love to spend eight hours talking to you about over several mojitos to be honest with you, but I just, I think we have a way of sort of helping the process now.

And I think we're moving with ONC on how we can facilitate this further.

CO-CHAIR SEPTIMUS: So, see, you missed the rooftop last night, Lillee. You should have come up to the rooftop. We would have had this conversation.

MR. GOLDWATER: I had a date night with

my wife. I'm sorry. The kids are gone for a week so I'm taking advantage of this.

CO-CHAIR SEPTIMUS: Shouldn't he be in radio?

MR. GOLDWATER: I was. I was in college. Can't you tell by my voice.

CO-CHAIR SEPTIMUS: Listen to that voice, that terrific voice.

MR. GOLDWATER: Helen loves that.

CO-CHAIR SEPTIMUS: We want to move forward because I want to make sure we get done. One more comment and then we're going to go forward. Yanling, go ahead. One more comment.

MEMBER YU: Thank you. Maybe somewhat related. Could you give me or help me understand what is the overall percentage of the facility or nationwide that adopted the EHR?

MR. GOLDWATER: So, the most recent data is that from Chilmark Research which is sort of this independent -- it's not affiliated with HiMMS, it's not affiliated with a lot of the vendor organizations.

I think that they said as of this year about 84 percent of all hospitals have EHRs and roughly 79 percent of all physician networks, physician centers have EHRs.

And what was even more interesting is that community health centers are almost over 90 percent adoption right now because of the funding that has been made available through organizations like HRSA that they've really become much more adept to incorporating EHRs, even than some large hospitals or physician networks.

But it's substantially higher than it was 15 years ago.

CO-CHAIR THRAEN: That is hospitals, not nursing homes?

MR. GOLDWATER: Yes, correct.

CO-CHAIR THRAEN: The whole continuum of care is not at the table yet.

MR. GOLDWATER: Health facilities are way behind. Long-term care, post-acute care is significantly behind. Ambulatory surgical centers. Long-term acute care hospitals.

Inpatient rehab facilities. They're much further behind.

MEMBER YU: Just one quick one. We have reviewed some records and we're trying to develop a policy, EHR policy for the medical board. And it's not lukewarm I can say for the physicians, for Washington State.

I just wondered down the road if will be created some type of incentive or something to help adapt this EHR for professional.

CO-CHAIR SEPTIMUS: There is Meaningful Use. There was incentive for physicians to adopt EHRs. So those incentives are actually there.

So I'm going to be forced to move forward.

MR. GOLDWATER: That's fine, Ed. Thank you very much.

CO-CHAIR SEPTIMUS: So if you'll introduce it, we're going to go to measure 3005, Initial Risk Assessment of Immobility-Related Pressure Ulcers within 24 Hours of PICU Admission.

This is from the Pediatric Consultants.

And this is an eMeasure. And it's a first time. So does this fit into the trial use? I don't think so.

DR. WOODS: No. We were able to test it in both Epic and Cerner which are the first and third in market.

CO-CHAIR SEPTIMUS: That's what I read, I just wanted to make sure. So this is a regular evaluation. So if you'll just give us a brief presentation then one of our group will lead the discussion.

DR. WOODS: Can I announce who's on the phone?

CO-CHAIR SEPTIMUS: Oh, please do. So who's on the phone for measure 3005?

DR. SACHDEVA: Hi, good morning. This is Dr. Ramesh Sachdeva and I served as the PI for the pediatric measurement center of excellence which was involved in the development of this measure.

CO-CHAIR SEPTIMUS: Is that it? Is

that all you're expecting? Is there someone else?

DR. WOODS: There should be a few others.

DR. MIKHAILOV: Hello?

CO-CHAIR SEPTIMUS: Yes, please tell us your name.

DR. MIKHAILOV: This is Theresa Mikhailov. I'm one of the pediatric specialists at Children's Hospital of Wisconsin and I was a member of the team that developed this measure.

MS. FOX: And I'm Jamie Fox, one of the critical nurse practitioners from Children's Hospital of Wisconsin.

CO-CHAIR SEPTIMUS: Thank you. So, is that CHOW versus CHOP? I'm kidding. Go for it.

MR. RICE: I'm Tom Rice. I'm the pediatric intensivist and I was the chair for the pediatric expert work group working on the PICU measure development.

DR. WOODS: Okay, so both of these measures were specified as eMeasures. All of our measures, actually, were specified by eMeasures,

ones that you're not seeing today.

Our experience is that pediatric hospital EHR systems are much more evolved than the ambulatory care context.

So we were able to test these in two different EHR systems, both Epic and Cerner. And I'll start by introducing the pressure ulcer measure.

Pressure ulcers develop when soft tissue is compressed between a bony prominence and an external surface for a prolonged period.

This results in tissue hypoxia causing cellular death, injury to the surrounding area and ultimately a pressure ulcer.

A pressure ulcer is a localized injury to the skin. Pressure ulcers have been steadily increasing with reported rates of 4.14 pressure ulcers per 1,000 pediatric discharges in 1999.

And it's up to 4.33 pressure ulcers in 1,000 pediatric discharges by 2002, and has increased 34.5 percent from 2000-2007.

Pediatric patients who experience

pressure ulcers have 6.15 percent mortality, and pressure ulcers can lead to infection, pain management challenges, disfigurement, increased length of stay and readmission, altered body image and psychological distress as well as considerable cost to the healthcare system.

Early intervention can be an effective prevention measure against pressure ulcer development.

Pressure ulcer prevention means an accurate assessment to identify at-risk patients. The Braden Q is the only validated immobility-related pressure ulcer risk assessment tool available for critically ill or injured children.

Identifying patients at risk for pressure ulcer and then intervening accordingly can reduce the incidence of these pressure ulcers which ultimately reduces infection, pain, disfigurement, length of stay, readmission, psychological distress and mortality in PICU patients.

The numerator for this measure is the number of PICU patients for whom an assessment of immobility-related pressure ulcer risk using a standardized pressure ulcer risk assessment tool was documented within 24 hours of admission.

The denominator is all patients admitted to the PICU for at least 24 hours during a monthly or quarterly reporting period.

The data source for this measure is the EHR as an eMeasure.

Performance scores. Children of all ages at risk for -- sorry, that's a different thing.

The performance scores, we were able to actually calculate the measure in both the Epic and Cerner systems, but we actually only had by the time of the testing the performance scores for the Epic calculation.

We had 100 percent reliability when the same set of charts was reviewed through manual chart abstraction.

And electronic output was provided for a reporting period of January 1 through March 31,

2015, included 106 unique patients representing 109 events.

We were able to output performance across different patient factors, age zero to 6 92 percent, 6 to 13 years 94 percent, 13 to 19 was 95 percent.

For race/ethnicity white patients had this assessment done within the 24 hours 97 percent of the time. For African-American patients 82 percent of the time. Hispanic 94 percent, and other 92.

We actually tested the eMeasure feasibility in five sites. It was found to be technically feasible in all sites. However, two of the sites dropped out for workflow issues.

So the structured field existed, but people didn't use it.

I think that gives you a good sense of what we've done.

CO-CHAIR SEPTIMUS: Are you ready, Pat, to take us through the evidence?

MEMBER QUIGLEY: Thank you, Mr.

Chairman, and thank you for the opportunity to present a second measure.

So with that my opening remarks, and thank you for the introduction to the measure, is to pose the question in terms of quality as this measure came to us.

This is a measure, a yes/no measure, of whether or not an assessment is done in the pediatric intensive care population.

And when I first read this, and I know others have commented in their reviews. When I first read this measure as it came forward to us it took me back to the days of the patient safety complication steering committee that was co-chaired by the current president of the American Nurses Association Dr. Pamela Cipriano.

And we discussed at that point in time what is the measures of quality. And we had very lengthy dialogue that a measure, yes or no, is something done or not, is not a measure of quality, unless it is really aligned into a composite measure.

And I take you back to a measure that came to us in relationship to fall risk assessment in the home care setting.

And when that came to us one of my comments then, and I know it would be on record, is that this measure here is nursing practice. It's nursing practice to assess patients across all settings of care, whether or not they're assessed for pressure ulcer risk.

So, just having an assessment done yes or no is not necessarily a measure of quality.

So, to me I really questioned how this came forward to the patient safety committee as a measure of quality.

But that being said, and I think it is a topic of discussion for this whole committee, is if there's one piece, if the intent is for this to become eventually a composite measure, then that's another discussion.

But to just measure nursing practice, is it being done or not, in a very at-risk patient population, pediatric, yes or no, isn't the

responsibility of the nursing profession.

So, this measure, when you look at the evidence for this, I was really surprised to not see a systematic literature review, and to have it be narrowed down to the pediatric population.

There's not a systematic literature review. There's not a grading of the evidence. The evidence that's presented to us is quite dated.

You could go through it and it's also cited in the preliminary analysis.

But as we look at the evidence to support yes of course risk assessment is essential before you do care planning because you have to have pressure ulcer risk to identify who's at risk for pressure ulcer.

And I would think in the pediatric ICU everybody's at risk for pressure ulcer.

But the literature that's presented to us is a 2001 survey, a 2006 guidelines for assessment of prevention in the pediatric population, a 1996 identification of skin integrity in the pediatric population, and 2003

review of pediatric care.

So essentially those are clinical guidelines. And the measure evidence was graded as moderate, but we all know that to prevent any outcome you always start with screening and get to assessment.

So that is why I really expected more. And I was really quite surprised to not see any comments by the Pediatric Nurses Association that were submitted. There were no public comments Pediatric Nurses Association.

So, I don't know if this went out to them. I know it's an electronic measure.

So, those are really my questions. My question, number one, is is it really a quality measure.

And then my other question truly is the amount of evidence to support it.

In the discussion of the evidence it indicates that there's really currently no clinical guidelines for this patient population.

But colleagues, two days ago the Agency

for Healthcare Research on Quality released its quality indicator toolkit. And that quality indicator toolkit had pressure ulcer prevention for the pediatric population.

It's the entire pediatric population. And all children should be assessed for pressure ulcer risk in 24 hours.

So, again, to be able to have that quality indicator come forward there has to be a body of evidence to support it. So those are my opening comments.

CO-CHAIR SEPTIMUS: Thank you. I want to just clarify one thing. So, is there anything that you reviewed, and the developer can also answer this, that by doing the assessment in the first 24 hours, that that is linked to a better outcome?

Because I think that's actually, and you guys correct me, that's really the standard for a process measure. And is there literature to connect those two?

MEMBER QUIGLEY: Yes, but what I'm

saying, what they presented is the data. But I know that there's more current. But we all know that that should be done. So that's why I was looking for.

DR. WOODS: So, just to put this in context. The measure developers, PMCoE, all the folks on the phone and me, are a center of excellence funded through AHRQ as part of the pediatric quality measures program that was hardline written into the CHIPRA law. Because there's a real paucity of pediatric measures.

There aren't PICU measures. These are the first PICU quality measures coming forward I'm pretty sure for PICU. No? Okay. Potentially the first eMeasures then.

So the focus on pediatrics is part of the program. And as you said -- so, I haven't seen the thing that you saw, but I know that AHRQ required us to present all of the information that you guys also have, and asked us to build fact sheets based on that, and have those fact sheets up on their website.

So I'm not sure if that's the material you're pointing to. It would be our stuff probably.

CO-CHAIR SEPTIMUS: Kendall.

MEMBER WEBB: So, as somebody who's pediatric trained I'm just going to provide the other side of this.

CO-CHAIR SEPTIMUS: Get a little closer to the microphone.

MEMBER WEBB: As somebody who's pediatric trained I just want to give an alternate thought about the evidence in this case.

It is notoriously hard to do studies on pediatric patients, almost impossible. Because if you create a situation where you show harm on one side almost everybody shuts the study down immediately because there are children involved. So you're going to have trouble finding high-grade evidence for almost any pediatric process, anything you try to put up here.

The one thing I would say is, you know, I'm not sure about the Braden Q although a quick

Google search of it, it shows up pretty heavily as the pediatric pressure ulcer risk assessment tool.

It is, as I can see, British, but it looks like a lot of people are using it.

While I agree that everybody should be getting pressure ulcer checks within the first 24 hours of arriving to an ICU setting it's clear even from what they did, they had two sites drop out because they felt like it was too much flow -- change. So it's clear that not everybody's doing it.

So, if we're really looking at patient safety, I'm not saying the evidence is there, I'm not saying it's high, anything like that, but if we're really looking at patient safety to me this does seem like a good place to start, especially with an eMeasure because it's a pretty easy eMeasure to get going.

CO-CHAIR SEPTIMUS: Missy.

MEMBER DANFORTH: Thank you. So, actually thanks to Dr. Steve Lawless I've had the opportunity over the past 12 months to have some

very significant conversations with Solutions for Patient Safety which is a pediatric collaborative out of Ohio.

They've put into operation an outcomes measure for pediatric patients throughout the hospital including ICU looking at incidences of pressure ulcers stage II and worse.

So can you just talk a little bit about the need for a process measure when it seems like there's a lot of pediatric hospitals that are looking at outcomes measures, and why you chose not to bring an outcomes measure forward?

DR. WOODS: This measure is to be a part of a set that would look at process and outcome. It's kind of the first in that process.

MEMBER DANFORTH: What's the rest of the set?

DR. WOODS: It would be outcome measures.

We tested in our measure champion hospitals a measure around I believe it was pressure ulcers grade III or higher.

And we assessed that as part of a composite measure of preventable harm. It includes other things like CAUTI and CLABSI. But we were able to assess the outcomes as well.

MEMBER MCGIFFERT: I'll just be very brief. I appreciate the focus on this population. It's a real problem.

But I think we need to see outcome measures. I'd like to see -- if you're developing some kind of composite I'd rather see that come forward than a check the box kind of measure.

DR. WOODS: Possibly some of my critical care colleagues could speak up here on the phone.

DR. SACHDEVA: Absolutely. If I may start here, this is Dr. Sachdeva. And I request my other colleagues to weigh in too.

So, I just want to make a couple of quick points. Besides the PI on this particular center where this work was performed I've also been a practicing pediatric critical care physician for several years.

And as you all know this is a clinically huge challenge for us in the PICU across the country.

And clinically I think going to an intensive care unit, most intensivists, physicians and nurses and other staff would agree that this is a clinical challenge which can put children at risk.

This is the first step of a longer journey. And as correctly pointed out previously by one of our colleagues in the room there is a paucity of pediatric evidence in general.

And I think the fundamental question which needs to be asked is how long do we wait clinically to obtain that necessary evidence before getting started.

This is not the end of the journey. This process measure is the first step of much more to come. But this is the beginning.

And my own feeling clinically is that not doing this, the potential risk posed to children is high.

And conversely I would also request the committee to consider that by doing this, if this measure were to be supported by the NQF are there any potential risks to children.

And I think most would argue that there are none. I mean, doing this doesn't solve the problem but is the first step toward solving the problem with relatively minimal burden if you may.

So, our measure already attempted to make sure that this is a process measure. This could be tested in EHR systems.

This would be an eMeasure which is another first.

So again, I think we need to look at it in light of the first step towards others.

But maybe Theresa with Jamie on the call can weigh in. You are content experts in this particular area.

DR. MIKHAILOV: This is Theresa. I was going to point out that, yes, was meant to be the first of a series of measures designed to address the problem of pressure ulcer prevention

in our critically ill children.

This is the first measure because it's the beginning of the process, looking at the skin at the time the patient arrives.

The second measure was meant to be an ongoing Braden Q in centers that are actively working to prevent pressure ulcers as Braden Q is done at admission but it's also done at least every 24 hours thereafter. So looking at that as a process measure was our second measure in the set.

The third measure, if there are problems identified with the Braden Q was to intervene with preventive measures for the patient. So that is also a process measure measuring whether the appropriate interventions or any interventions in fact are made to prevent development of pressure ulcers.

The fourth was the outcome measure looking at the rate of pressure ulcer incidence but from immobility-related. But there is now a surge in device-related pressure ulcers as well.

And the Braden Q is not as well designed

to address that. So, we were intending that to be the fifth measure.

But we feel that the process measure has to precede the outcome measure. If we only look at the outcome, the outcome will be worse. If we look at the process measure we think we will have a positive impact on the outcome by intervening before the outcome occurs.

DR. WOODS: And that was Dr. Theresa Mikhailov who's a pediatric intensive care clinician.

CO-CHAIR SEPTIMUS: So let's get these last few comments. I'd like for us to get to the meat of the issue and find out whether the committee feels that the evidence is there to move forward.

DR. WOODS: Could I say one more thing?

CO-CHAIR SEPTIMUS: Of course.

DR. WOODS: There -- it's been very difficult to get measures implemented into the Medicare/Medicaid program which was part of the intent.

This pediatric quality measure program

was built into the CHIPRA law so that CMS could have better data and the states could have better data. Many states are not implementing these measures.

But California is very interested in evaluating both of the eMeasures that are here. The head of California as a part of another proposal, sort of the second round of the PQMP program has already signed onto these measures.

CO-CHAIR SEPTIMUS: So three more comments and then I'd like to go to the vote. So Steve, Iona and Yanling.

MEMBER LAWLESS: This is Steve Lawless. I'm a pediatric intensivist, 30 years.

To Pat's point this is basic nursing assessment. I mean, this is really what it is.

I think if you look at validity, reliability, intent, no argument. I mean, this is what you have to do.

However, the Braden score is 28 elements. It's a lot. And so, it's a lot. And I like what the developer said in terms of this is a stepped approach.

And I think on the time of admission within 24 hours, and we heard yesterday how long it takes for a pressure ulcer to form, you're getting a baseline versus progression.

So, I think the idea is all the work is going to be involved. When is this ready for prime time in development.

I would argue that you may want to consider having a lot more maturity to even -- because you're putting all this stuff in.

And I think feasibility and usability is going to be a big issue here. The Braden Q, it's a good scoring system, it is a lot of data. And you have to balance that if I'm moving patients around and everything else, and what I'm doing with each of these 28 elements. So it's a lot of stuff there.

Not to argue against the importance of it, but this is bigger of a workload than what's coming across here.

And you are diverting nurses from bedside care to filling out a scoring system of 28

data elements. So that piece, you have to see what's that impact.

DR. WOODS: Jamie, do you want to speak up here? Our pediatric intensive nurse.

MS. FOX: I think it depends on what the hospitals are doing now. At least here at our institution the nurses are pretty familiar with doing this. So it isn't as labor-intensive as it would be to implement it as a brand new tool.

It's built into our Epic system so it's part of their standard questions that they ask and fill out as their assessment.

CO-CHAIR SEPTIMUS: Iona.

CO-CHAIR THRAEN: So, this is a question for NQF staff.

Is it -- could the committee recommend that this be not endorsed but put into the trial option? If the committee feels that it is not quite ready, or they want it to be bundled with the outcome measures in the future but there needs to be some testing that takes place, et cetera, et cetera. Is that an option?

CO-CHAIR SEPTIMUS: Before I let them answer, in the back of my mind I thought we had considered an outcome measure for pediatrics around pressure ulcers.

And our great support folks here pulled up actually something we approved that was an outcome measure in pediatric for pressure ulcer. So I should have gone back and looked myself, but you may want to comment on that.

MR. LYZENGA: So let me just first address the trial use.

I think -- I'm not as familiar with the policy, but my understanding is it's usually we want a measure to sort of pass the other criteria, evidence, importance, before -- it has to -- before we can put it into trial use. That is really where the -- for where the testing hasn't been done, but it has met all the other criteria.

So I think if we're hesitant on evidence here we would want to pass it on evidence before we could consider the trial approval.

But we did --

CO-CHAIR THRAEN: Then for follow-up to the measure developers one of the observations that were made earlier was that the evidence that was provided is pretty dated.

And if you're a center of excellence and you're involved with this process can you tell us why that evidence is so dated?

DR. WOODS: Dr. Mikhailov, can you respond to that?

DR. MIKHAILOV: I think that there is some that is more recent, and there is something that I was just informed of yesterday in this area that was just accepted for publication yesterday by one of my colleagues.

So, I couldn't add that obviously. It's in press probably today, so I can't share that with you, but I can tell you that it supports that this is a critical issue.

These are patients that were followed with -- in our institution with Braden Q at admission, Braden Q every day, and a two-year cohort of patients of whom 19 developed severe

pressure ulcer unstageable types or deep tissue which are the worst type.

And 42 percent of the 19 patients passed away with these severe. And these were all patients who were having all of these interventions that we've outlined in our series of measures. These were all immobility-related pressure ulcers as well.

So I know that there is literature coming out as Dr. Sachdeva told you. These are difficult studies to do. So that was a retrospective review of existing patients.

There isn't a good prospective study.

CO-CHAIR SEPTIMUS: Thank you very much. Yanling, one more comment and then we're going to vote.

MEMBER YU: Okay. My understanding is that for this type of a check on yes I documented, no I didn't, basically it doesn't really relate it to the outcome until you check whether you documented and not documented shows any differences in the outcome down the road.

So, I'm assuming the idea is to collect the data down the road and to better understand the outcome as you said, the plan.

So now my question is with the current medical record system do you have any idea when you look at others that -- whether the -- what I'm trying to say -- how much improvement would be after you use as an eMeasure to really get a better outcome, or a better documentation of whether the children's pressure ulcer have been prevented after you do this type of documentation. Do you know my question?

DR. WOODS: I'm not entirely sure I understand your question so I'll say it back to you and see if this is what you're asking me.

If this is really only about documentation versus about doing the assessment.

So, in one of our sites -- so, all of our sites were doing some of the Braden Q because we also did chart reviews, and we thought that presenting an eMeasure was a less burdensome activity.

What some of the systems, the way the workflow happened they would do it on a paper form that was then scanned in.

And as you heard the person who was sitting next to me, you can't get that information. It's not a structured field.

So, when they document it in electronic records they had the fields but their process was to -- when they did the Braden Q their process was to scan a document in as opposed to note each of the elements in the electronic record.

So, all of the elements when they did it were in the electronic record, but just in a scanned document.

So it's an easier measure -- if you're going to do the right thing which is to do a pressure ulcer assessment it can be done on paper or it can be done in electronic fields.

And if the EHR in both Cerner and Epic had electronic fields for those elements of the assessment and then those were used and then were able to be used for construction. Is that

answering your question?

CO-CHAIR SEPTIMUS: I think so. I think I know where you both are going. But Laura has a quick question and then we really do need to vote.

Let's vote.

DR. MIKHAILOV: Can I make one quick comment? This is Theresa Mikhailov again.

The Braden Q has a maximum of 28 points, but it's in 7 fields. So it's not 28 separate fields that are entered, it's 7 fields with scores in each field. So I think it's not as burdensome as it might seem.

MR. LYZENGA: It's 28 decision points.

CO-CHAIR SEPTIMUS: Okay, let's go ahead and vote.

MS. QUINNONEZ: We are now voting on measure 3005 initial risk assessment for immobility-related pressure ulcer within 24 hours of PICU admission. Voting is now open for evidence.

Your options are option 2 moderate,

option 3 low, option 4 insufficient. Those voting options are option 2 moderate, option 3 low, option 4 insufficient.

Michelle, if you could submit your vote in the chat box, please?

MEMBER SCHREIBER: I did. Did it not go through?

MS. QUINNONEZ: Okay. All votes are in. Voting is now closed.

The votes for evidence of measure 3005 are 32 percent voted moderate, 47 percent voted low, and 21 percent voted insufficient.

CO-CHAIR SEPTIMUS: Okay, well that is a no so I think we stop.

MR. LYZENGA: Evidence is a must-pass criteria.

CO-CHAIR SEPTIMUS: Evidence is a must-pass.

CO-CHAIR THRAEN: I have to get a clarification question. I'm still struggling here a bit.

So, if the evidence -- so in the

instance of eMeasures which are the new kids on the block, if there is no evidence because you haven't used an eMeasure to determine whether in fact it --

(Simultaneous speaking)

CO-CHAIR THRAEN: So you're just asking about the content of the measure.

DR. BURSTIN: It's evidence for the measure focus, not as applied.

CO-CHAIR THRAEN: Okay. All right.

CO-CHAIR SEPTIMUS: But really, thank you very much. I know it seems a little painful.

We're going to go ahead and go onto the next measure. And I think you're going to also be 3006?

DR. WOODS: Yes.

CO-CHAIR SEPTIMUS: Let me introduce the measure first. Initial Baseline Screen of Nutritional Status for Every Patient within 24 Hours of PICU Admission.

And our measure developers will make a few comments. This is also an eMeasure and a

process measure.

DR. WOODS: Right. Developed by the same group. And it is a process measure. So I'm hearing that process measures are less interesting to this group.

DR. BURSTIN: NQF has a stated preference for outcome measures. It's in everything we do. So it's not this group, it's actually NQF-wide.

DR. WOODS: Okay.

DR. BURSTIN: We prefer outcomes. If they're process measures, they have to have a clear evidence link to outcomes.

DR. WOODS: Okay. In critically ill children malnutrition is associated with an increased PICU length of stay and an increased risk-adjusted mortality rate.

Identifying nutritionally at-risk patients as early as possible in their illness allows providers to prescribe nutrition therapy that is appropriate for patients' nutritional status and clinical condition that will most

effectively facilitate the healing process.

In an initial baseline screen nutritional status for every patient increased awareness of the patient's nutritional state, specifically identified the subset of PICU patients who are at risk of malnutrition, and allows providers to adjust the timing, content, quantity of nutrition therapy to meet the individual patient's needs.

While there is no single validated screening tool, institution-derived nutrition screening tools can be used, typically take about five minutes to administer, can be performed at the bedside and do not generally involve a dietitian.

Screening of nutrition status is fairly quick yet vitally important as the benefits of nutrition support in the critically ill patient include improved wound healing, decreased catabolic response to injury, improved gastrointestinal structure and function, decreased PICU length of stay and decreased mortality.

It is based on a clinical guideline of 2009 where it is stated that children admitted with critical illnesses should undergo nutrition screening to identify those with existing malnutrition, or those who are nutritionally at risk.

The specifications of the measure. The numerator is the number of PICU patients for whom a screening of nutritional status was documented with use of a standardized nutrition screening tool within 24 hours of admission to the PICU.

The denominator statement is all patients admitted to the PICU for at least 24 hours during a monthly or quarterly reporting period.

And there's a denominator exclusion of patients who have already had a documented nutritional screening or assessment in the previous 48 hours.

The data source is the electronic medical record which constructs the measure as an eMeasure.

The measure was tested in four different hospital systems, two EHR systems, Cerner and Epic. Electronic output was provided for 110 unique patients representing 121 events.

Clinical performance represented by the results of the eMeasure was 90 percent of patients and 92 percent of screens meeting the measure.

It was feasible in three of four institutions when feasibility was assessed. Again, this notion of a scanned-in document was what the workflow issue in one of the institutions.

One hundred percent -- when reliability was assessed on a set of the same patients' medical records through manual chart abstraction reliability was 100 percent.

CO-CHAIR SEPTIMUS: Can you perhaps just so some of the questions we had last time, can you give us the relationship of this measure to outcomes? How strong is that relationship? So just to get that out of the way now.

DR. WOODS: It's part of a guideline so

that's a bit more. Maybe Dr. Mikhailov or Dr. Sachdeva, would you like to comment on that?

DR. SACHDEVA: Yes, this is Rames. Maybe I'll refer to Theresa given that this is, again, an area for clinical and epidemiological expertise. Theresa, please.

DR. MIKHAILOV: So, again, this is a measure that was intended to be part of a series of measures. And these were all intended to improve nutritional status of patients in the pediatric ICU.

This was to be the first measure with nutritional screening at the time of admission. For those malnourished patients or at-risk the next measure would have been assessment.

And assessment is very different from a screen. A screen as you heard is something that should be able to be done in five minutes at the bedside by the bedside provider.

An assessment requires someone with a different skill set, usually a dietitian or a specialist in nutrition. That would be something

that should be done for those at-risk patients to intervene earlier.

And the third measure that was intended to be included with this was identification of caloric goals for the patient within 48 hours of admission.

All of these are measures that were meant to be done in sequence. The first one was the only one that came through in our wave one of measures.

CO-CHAIR SEPTIMUS: Okay.

CO-CHAIR THRAEN: So, I'm sorry. Thank you for the context, but what we're asking is what's the evidence that supports the assessment linking to nutritional outcome and patient harm in the pediatric population.

DR. MIKHAILOV: Well, I think there's an abundance of literature which I think we have included in here that malnutrition is very common in pediatric ICU patients.

It also is something that has been found to develop in a large proportion of critically ill

children. And we also know that children who are malnourished have a higher mortality.

There isn't a specific link between the screen because as opposed to the pressure ulcer scenario screening is much more diverse. There is no single validated tool that is used broadly. Many institutions use their own individually designed screen.

There are a handful of screens that have been validated in certain populations within children.

There is no one tool that has been validated in a general critically ill pediatric population.

CO-CHAIR SEPTIMUS: Okay, thank you. I guess we have two quick comments. Oh, excuse me, three, before we get to the evidence. Albert?

MEMBER WU: Yes. So, I can certainly see the importance of this issue.

I think that in proposing a process measure you are suggesting that perhaps we should be changing the way that we practice pediatric ICU.

And you are making a specific suggestion about how the nutritional screening should be done. And that it be done by someone with that specific background.

So as I look at this I'm thinking about our ICU and I'm wondering is this the way we should change practice. Is this the way we should all change practice.

Is there someone else who could do this -- is there a person of a different job description, is there a different method that could be used as opposed to this one, and should we absent evidence that this is either the best way or a way that is linked to those nutritional outcomes, should we be prescribing that at this committee?

DR. WOODS: Just to be clear, we're not recommending any particular tool. There are a lot of different tools out there. There are a lot of institutionally developed tools.

What we are recommending which is considered appropriate practice is that there be a screen done.

MR. RICE: The other point is that the screen can be done by any bedside caregiver, whether it be nurse, physician, or whatever. The screen is very quick and simple.

The assessment, however, for those that fall out and are determined to be malnourished would then move up to the -- usually it's a nutritionist or dietitian.

So that is not this measure. That was the other series of measures that Dr. Mikhailov referred to.

CO-CHAIR SEPTIMUS: Steve.

MEMBER LAWLESS: Yes, Steve Lawless. How does this differ from the Joint Commission requirement that everybody within 24 hours gets a nutritional screen or assessment?

DR. MIKHAILOV: This is actually based on that concept, but there is a somewhat nebulous definition of what is required by the Joint Commission and not all institutions meet that standard in the same way.

That is, however, why there are these

many institutionally derived screening tools without one or even several validated tools for pediatrics or anybody else.

MEMBER LAWLESS: Right and I get that. So the idea is you're not recommending a tool, you're recommending a screen.

DR. MIKHAILOV: We can't.

MEMBER LAWLESS: Right. But I'm just asking what's the difference then -- what is people are supposed to be doing anyway.

DR. WOODS: So, we don't have the data here. We also assessed this as a chart review measure and had performance scores for the chart review.

And in one of the children's hospitals in Chicago only 23 percent met the measure. So, we were pretty surprised by that. Only 23 percent met the measure in a children's hospital in Chicago.

But we were asked to take all of our chart reviewed data out of this for those that didn't meet the eMeasure. But I know the answer.

CO-CHAIR SEPTIMUS: Laura, then Pat.

MEMBER ARDIZZONE: Just one comment. It is very concerning to me that there's no validated tool. What are we measuring then? It sounds to me a little garbage in, garbage out.

If you're just making people do a measurement which may not be measuring anything reliable or valuable what is the point of making people measure?

Until you have a valid, reliable tool that can be implemented across the United States and mean something. Right now you're just having them measure nothing, really.

DR. WOODS: Dr. Mikhailov, might you respond to that?

DR. MIKHAILOV: Well, but the Joint Commission already requires that each institution has a means for screening. It did not require that there is a validated screening tool and so none currently exist to our knowledge.

There is an assessment tool that has been validated in critically children under the age

of 5, that's the SGNA, Subjective Global Nutrition Assessment. That was validated in our institution.

But it is an assessment tool and it can take as much as 30 minutes for an individual patient. And it takes a skilled provider which in our institution is a registered dietitian.

That's not really a feasible mechanism across the board for all children. And so institutions use their own individually derived screens.

We are here working on developing a validated screening tool, but it takes some time. We've been working on it for over a year. But it doesn't exist yet.

But the fact that screening has to occur is what we're trying to make sure happens. We know that institutions don't meet that. The 23 percent was not a surprise to us.

Our institution was cited in the recent past for not meeting the screening criteria as well. That is part of what prompted us to pursue

this.

We've also done research here looking into the effects of enteral nutritional and parenteral nutrition. And so we understand the relationship between nutrition and outcome is important. We don't have prospective data for that at this point.

But we think that this is a sequence that matters, identifying nutritionally at-risk patients, assessing the appropriate patients, intervening appropriately and then hopefully altering outcomes.

MEMBER QUIGLEY: Thank you. And again as a committee member -- this is Pat Quigley's voice -- I'd like to say that a screen measure is not an indicator of quality.

And in the AHRQ toolkit again that just came out two days ago the measure that -- what they are advocating for is actually daily rounds assessment of nutrition. It's much more than screens. So they have been much more articulate in what needs to be done in terms of hydration and

assessing nutrition in the pediatric population. Thank you.

CO-CHAIR SEPTIMUS: Yanling.

MEMBER YU: Yes. The question is regarding the recommendation for that particular -- apologize, I forgot the name, the evaluation tool to assess the risk of a pressure ulcer.

My question is different hospitals may adapt different tools. And some of them may find others may be more useful.

Do you have a plan, any thoughts on how do you -- to look at the difference, how you reconcile the differences when people or facilities use different tools to do the evaluation? Or is that an issue at all?

DR. WOODS: So, I think the idea is that a screen be done so that at-risk children can be identified, and then a standardized validated tool would be used to assess their nutritional status.

But if they're not being flagged they don't get assessed and therefore there's poor care and severe risk of mortality to these children

because they're not assessed.

They're not screened first to then apply the assessment.

MEMBER YU: Okay, but my question is not whether it's assessed or not. It's assessed using which tools. Because if I understand correctly you said that there are other tools may be available for them to do the assessment. Is that correct?

DR. WOODS: There are currently no validated -- oh, for assessment?

MEMBER YU: Yes.

DR. WOODS: Dr. Mikhailov, can you address that?

DR. MIKHAILOV: So, I'm only familiar with the SGNA assessment. I also know that it is not widely used because it is time-consuming.

Dietitians have other mechanisms of assessing a patient which are not necessarily a structured assessment tool.

So I think there is variation both in the screening and the assessment process. That

doesn't mean that those are wrong.

I think that screening is probably fairly similar even though the tools have different components. They are really designed to identify patients who are nutritionally at risk or who are malnourished versus patients who are neither.

And the assessment is only for the patients who are at-risk or malnourished, generally.

MEMBER WU: So, just to follow on that. So, you said, well, if the children are not flagged then they will not get appropriate treatment.

So, my question is how are they being flagged? Are they being flagged in a way that is valid?

If you flag someone who does not need something then obviously there will be no ill consequences of not following up. I need some evidence that some or any of these screening tools are in fact predictive of being nutritionally deficient.

DR. MIKHAILOV: I don't think that

exists because, again, there are only a handful of screens that have been validated in limited pediatric populations and they have not been used in that manner that you refer to.

MEMBER WU: So we are recommending that we use something which may or may not be valid and recommend that this be done nationally.

CO-CHAIR SEPTIMUS: Thank you, Albert. Missy?

DR. MIKHAILOV: It's already a requirement that they be done.

MEMBER DANFORTH: Yes, so just a couple of things to point out.

So one is that this measure, the evidence for this measure is a little different than the previous measure. For this measure they did submit a systematic review and the NQF staff actually graded it as moderate.

Yesterday we reviewed a med rec measure for dialysis center where the measure developers said there is no evidence that med rec alone will have an impact on reducing medication-related

errors. And that measure passed evidence when we voted the second time.

This is a measure where it's a high-risk population. There's no outcome measure in place. So it's not like the pressure ulcer where there are pressure ulcer outcome measures in place.

There's no outcome measures associated with this particular measure. They're saying it's a first step. There is a guideline that supports it.

They submitted evidence that NQF staff graded as moderate. I'm just bringing that up because I think it's important that we grade these evidence things consistently and we look consistently at what's been submitted.

And there's a lot of parallels I think between this measure and the measure we looked at yesterday where the developer stated in this room there was no evidence that med rec alone had any impact on outcome, that it was a first step.

This measure developers is saying there's a guideline in place. There's a Joint

Commission core measure in place to do a screen, and it's measuring whether or not the screen is happening, and that they're going to be bringing additional measures forward related to assessment.

So I just want to bring that up.

CO-CHAIR SEPTIMUS: Thank you, Missy. I think we've actually talked a lot about the evidence so I think we can go to a vote on the evidence.

Although we haven't gotten to the gap yet there does clearly seem to be a gap. And I think the question for the committee is given there's some variability in screening, and how that's assessed, and whether or not there's an action taken on that screen that affects outcomes I think you're going to have to weigh the evidence that they presented and decide whether the evidence is strong enough to go onto the CAC. So, that's how I sum it up.

So, why don't we go ahead and go to the vote on evidence.

MS. QUINNONEZ: We are now voting on

measure 3006 Initial Baseline Screen of Nutritional Status for Every Patient within 24 Hours of PICU Admission.

Voting is now open for evidence. Option number 2 moderate, option number 3 low, option number 4, insufficient.

MEMBER DANFORTH: They submitted a systematic review. We should be able to vote 1.

MS. QUINNONEZ: Sorry, here we go. We'll revote again. Sorry.

Option number 1, high. Option number 2, moderate. Option number 3, low. Option number 4, insufficient.

Option number 1, high. Option number 2, moderate. Option number 3, low. Option number 4, insufficient.

All votes are in and voting is now closed. Evidence on measure 3006 reads 10 percent high, 40 percent moderate, 35 percent low, and 15 percent insufficient.

CO-CHAIR SEPTIMUS: Okay.

MS. QUINNONEZ: Consensus not reached.

CO-CHAIR SEPTIMUS: But it's not that we can move forward with the other. So we'll move forward.

And go next to gap. I think we've already -- if you want to have more discussion we can, but it sounds like the developer has already presented evidence that there is a huge gap in screening.

Yanling, I don't know if you want to say anything else about the measure or anybody else. I think this one. Yes, Lisa.

MEMBER MCGIFFERT: So, am I reading this right that there isn't much of a gap between the eMeasures? But you indicated that there was a gap in the others which we are not -- in like chart reviews that we're not considering.

DR. WOODS: Right. In the three hospitals that had the capability both technical and through workflow to present zero to 6, 92 percent, 6 to 13 years old, 94 percent, and 13 to 19 was 95.

But we see a gap just for those eMeasure

sites. White had a 97.5 percent meeting the measure. Black had 82 percent meeting the measure.

For English-speaking 95 percent met the measure. Spanish, 88 percent met the measure.

So there are disparities demonstrated here, but also in our chart reviews we found more variability.

MEMBER MCGIFFERT: Okay, so there were disparities based on race, but not statistically significant, right? Is that what that says?

DR. WOODS: I believe they were statistically significant.

MEMBER MCGIFFERT: It says these differences were not statistically significant.

DR. WOODS: Okay. I have a thing that says that they were.

MEMBER MCGIFFERT: And I guess my other concern is --

DR. WOODS: Lindsay, if you could weigh in here. I thought they were statistically significant.

MEMBER MCGIFFERT: My other concern is that it looks like -- although we're hearing about the chart reviews this is an eMeasure and there are not going to be chart reviews.

It looks like we're pretty close to being topped out on these measures. Am I reading that wrong?

DR. WOODS: No, because we're looking at pediatric ICU care and it was technically feasible in the two other institutions, it just -- their workflow was to have a scanned document.

So we didn't have time in our testing institutions to make those workflow changes because of the mechanism of funding for this research. But they are willing to make that change.

So it is not just about -- I mean it was technically and workflow feasible in both Cerner and Epic, and it also is technically feasible and requires just a workflow change, clinical documentation change in the other institution.

MR. LYZENGA: Were you saying in those

institutions that couldn't do the eMeasure but did the chart review there was a larger performance gap is what you're saying?

MEMBER MCGIFFERT: And we don't really know if that performance gap was because of the way the measure was implemented with chart review rather than an eMeasure.

DR. WOODS: Same specifications.

MEMBER MCGIFFERT: And I just want to reiterate the close to topping out issue with a process measure.

CO-CHAIR SEPTIMUS: So, if I understand the eMeasure performed much better than the paper.

DR. WOODS: The sites who could implement an eMeasure performed better.

CO-CHAIR SEPTIMUS: But you found a gap in a non-eMeasure site.

DR. WOODS: Found a gap in disparities in all sites. And we found a greater performance gap in sites that could not implement the eMeasure because of workflow issues, not because of the

missing elements.

CO-CHAIR SEPTIMUS: Okay, Pat?

MEMBER QUIGLEY: Thank you. My question goes back to the evidence. Did we pass the grade for 60 percent and higher to continue?

MR. LYZENGA: We hit 50 percent. So it was consensus not reached.

MEMBER QUIGLEY: Oh. Thank you.

MR. LYZENGA: So we do move on. We'll have to revisit that.

MEMBER QUIGLEY: Thank you so much.

CO-CHAIR SEPTIMUS: Albert? Any other comments before we vote on gap? Okay, seeing none we'll vote.

MS. QUINNONEZ: Voting is now open for performance gap of measure 3006. Option number 1, high. Option number 2, moderate. Option number 3, low. Option number 4, insufficient.

All votes are in and voting is now closed. For the performance gap of measure 3006 zero percent voted high, 50 percent voted moderate,

45 percent voted low, and 5 percent voted insufficient.

MR. LYZENGA: So again we're in the gray zone there, consensus not reached, but we'll move onto the next criterion.

CO-CHAIR SEPTIMUS: Reliability.

CO-CHAIR THRAEN: Just to clarify again. So, if it's between 50 and 60 percent we're considered in the gray? Forty and sixty percent. If we have 60 percent we have consensus.

So I get confused when we stop versus moving forward even though we haven't reached consensus. What's the difference?

So, there was a measure that we stopped on.

MR. LYZENGA: That was because we had greater than 60 percent voting against.

CO-CHAIR THRAEN: Against it. That's what it is. All right, thank you.

CO-CHAIR SEPTIMUS: Yanling, can you talk about reliability? Because Linda's not here and I asked Yanling at the last minute to do this.

So we thank her for that.

MEMBER YU: The reliability test. The eMeasure test was conducted in four Chicago area hospitals. It was only conducted in one of those hospitals because implementation issue at other three.

To demonstrate reliability the developer did element testing at one hospital site with 288 pediatric beds including 40 PICU beds and approximately about 11,291 pediatric admissions annually.

The testing period is between January 1 till March 31, 2015, at the one children's hospital.

And the analysis had 105 unique patients representing 121 events. That's something I don't understand, 105 unique patients but have more event numbers than that.

The testing involved eMeasure, also a computer score automatically, a manual chart review.

The results for testing shows that the

inter-rater reliability was conducted on five patient charts. Agreement is 100 percent for all critical data elements and 100 percent for overall clinic performance.

Because the agreement is 100 percent a cover score could not be computed.

So, that left the preliminary reading for reliability is moderate.

So the question for the committee is is this best sample adequate to generalize for widespread implementation.

And the second question is do the results demonstrate sufficient reliability so that different performance can be identified. So those are the questions suggested to the committee.

CO-CHAIR SEPTIMUS: Any comments from the developer or from the committee on reliability? Okay, we'll vote.

MS. QUINNONEZ: Voting is now open for the reliability of measure 3006.

Option number 1, moderate. Option number 2, low. Option number 3, insufficient.

Option number 1, moderate. Option number 2, low. Option number 3, insufficient.

MR. LYZENGA: Just to clarify moderate is the highest potential rating because it was -- testing was only done at the data element level, not at the measure score.

DR. WOODS: No, we also had the performance score. I can show you. I just read off of it and it is -- the patients that were 19 plus years old, there was a significant difference between those and others. I can show you. I mean, this is from the eMeasures, not from the other data.

MR. LYZENGA: What kind of analysis did you do on the measure score?

DR. WOODS: We calculated the measure. We calculated the measure and we also conducted reliability testing with a manual chart abstraction gold standard.

MR. LYZENGA: We consider that data element reliability, not reliability of the measure score.

DR. WOODS: And we also demonstrate the

difference in performance across different patient populations, age, race, ethnicity, language, and insurance status.

MR. LYZENGA: So that's more performance gap.

For reliability of the measure score we're looking for something like a signal to noise analysis of the measure's ability to differentiate between scores of different facilities.

DR. WOODS: You are only accepting signal to noise these days?

MR. LYZENGA: Well, we accept data element reliability as well, but for the measure score testing we want to see that the measure -- we want to see that you've -- that the measure score can reliably distinguish across different facilities.

So we wouldn't be able to do it just at one single --

DR. WOODS: So I can answer that. We have more than one facility. But the funding mechanism wouldn't allow us to have them give it

to us during that period. We'll have another round.

If you want that data in a few months, or like even a month.

CO-CHAIR SEPTIMUS: No, no, that's not your fault, but based on what Andrew just said because it's only done in one facility we can't look at multiple facility reliability. That's the signal to noise that we're talking about. So therefore we can't consider it as a high level.

Okay, so why don't we start the voting process all over again.

MS. QUINNONEZ: We're good. We haven't totaled yet. All votes are in.

CO-CHAIR SEPTIMUS: Oh, okay, never mind.

MS. QUINNONEZ: All votes are in. Voting is now closed.

CO-CHAIR SEPTIMUS: Let me ask you, was there confusion about this before you voted? Because we can easily vote again.

MS. QUINNONEZ: We'll revote.

CO-CHAIR SEPTIMUS: Let's revote just to make sure, okay? This is on the reliability.

MS. QUINNONEZ: This is for the reliability of measure 3006. Voting is now open for the reliability of measure 3006.

We're looking for one more vote. Can you resubmit your votes, please.

Voting is now closed. For the reliability of measure 3006 37 percent voted moderate, 42 percent voted low, and 21 percent voted insufficient.

MR. LYZENGA: So that --

CO-CHAIR SEPTIMUS: -- must-pass measure.

MR. LYZENGA: So that means that the measure does not pass on reliability because it's the two low and insufficient together gets us over 60 percent.

DR. WOODS: As I mentioned the funding mechanism for this, we were not allowed to do anything after the date of its end. But it was already done, they just couldn't pass the data over

to us. It's kind of an odd thing.

If I can submit that data that shows across three different institutions?

CO-CHAIR SEPTIMUS: That would certainly help. Believe me, I think we all feel the same empathy. I think you've really done a great job with what you had.

This is a rather rigorous process and I think that when you get those other data elements and other things in place I think you'll get a different reading from the committee.

But I don't want you to walk away feeling unwanted. We do want you to come back. And I think now that you see -- I really feel bad when we turn down measures, but I want the developers to realize that we know you did your best job, and there's just certain things.

I think now from the discussion, I think you now know that we need to pass a measure. I have no doubt once you get that additional stuff in place you can bring that measure back to the next round and you'll get I think a different read.

In another set of NQF meetings around measures there was a similar kind of problem, that some of the data just wasn't able to be passed forward. And there was a phone call or something. We were allowed to --

MS. MUNTHALI: Yes. So you can bring it back during the post comment call. So, during this commenting period you said you'd have -- you could do it within a month? So you'd have about two months to do that.

DR. WOODS: Okay, great. Thank you very much.

CO-CHAIR SEPTIMUS: Thank you so much. You really did a very nice job. Hopefully we can -- of course you can say so.

MEMBER MCGIFFERT: Thank you for bringing this. And I think -- I couldn't catch everything that Ed said, but I think it's really -- I mean you've heard some of our questions.

And I would just say I know I would look more favorably on it if there were some outcome components with it. So I hope that you guys will

work on that. And it sounds like you are working on it and will bring it back to us with a little bit further measure. Thank you.

CO-CHAIR SEPTIMUS: Thank you very much. So we're going back to 2983 Potassium Sampling Hemolysis in EDs. The Cleveland Clinic is the developer so are they on the phone or are they in person? They're on the phone?

DR. PHELAN: Yes, sorry, I missed the last part. I was getting off mute. This is Michael Phelan speaking.

CO-CHAIR SEPTIMUS: Fine. Can you tell us who you are? Are you in Cleveland?

DR. PHELAN: Yes. Believe-land as we say now.

CO-CHAIR SEPTIMUS: Well let me tell you, Cleveland ended up very much better in terms of outside protests and stuff than Philadelphia has.

DR. PHELAN: This is true. And the Cavs won the national championship in basketball so we're in Believe-land now. So yes, I think it

made us look very positively. I thought it went very well, the convention and the national championship in basketball.

CO-CHAIR SEPTIMUS: Let's hope your measure does as well as the Cleveland Cavaliers.

DR. PHELAN: You set me up, man, this is too much.

CO-CHAIR SEPTIMUS: Anyway, if you could give us a brief overview of your measure and then we'll go through our usual discussion of the evidence, the gap, the reliability, et cetera. So, please.

DR. PHELAN: Sure. I am an ED physician and I always like to say I'm the accidental measure developers because I keep coming up with measures that I wasn't really expecting to come up with.

This measure came about through one of my colleagues in lab medicine was offered an ability to look at this from the CDC.

And he wanted nothing to do with it because he knew some of the challenges and the

struggles associated with this measure.

But hospital EDs have for years been identified as the major source of hemolyzed lab samples for hospital labs.

It's the leading cause of unsuitability of specimens. They rate significantly elevated compared to other departments within the hospital. The closest we have is maybe the ICU which is about half at least at our institution hemolysis rates.

In the literature ED hemolysis rates range anywhere from 6.8 to 30 percent. I actually found one paper now, a newer one that rated one at 67 percent.

The American Society of Clinical Pathology has a 2 percent lower hemolysis benchmark, but I can tell you from the quality people in lab medicine here at the clinic, their expectation from the people who work for them which are the phlebotomists, their expectation is less than 1 percent.

When blood samples are hemolyzed there's interference in over 39 different lab

tests.

The unreliable lab tests, especially potassium, and that's the one that we measured, but also they're rejected for coagulation studies and type and screens.

And falsely elevated potassium may indicate life-threatening abnormality, and an apparently normal potassium due to hemolysis may be hiding a significantly low potassium.

There are two groups that have identified this as a significant issue. First of all, the CDC, and I get back to how I got involved in this.

The CDC through a laboratory medicine best practices and systemic review meta analysis that's authored by Heyer and cited throughout here identified ED hemolysis as a significant problem in lab medicine.

And they funded a cooperative agreement to study this, and that's where we studied it cooperatively with lab medicine and nursing at our institution.

Like I said, it's such a significant problem that the CDC did this laboratory medicine best practices. They did a comprehensive search from 1990 to 2011. They found some 600 publications and abstracts, and 22 EDs and some non-published data.

The experts saw practice impact on hemolysis and identified seven practices within that.

Of the seven practices they could really only find two of them which they could recommend best practice at the time, and that was straight stick needle and antecubital location if you're drawing it through an IV.

They also identified four other things that may contribute but didn't have sufficient evidence which was syringe versus vacuum for the location, large gauge versus smaller gauge, partial vacuum tubes and tourniquet time.

We at the Cleveland Clinic started looking at this and we even did an analysis from the National Hospital Ambulatory Medical Care

Survey from 2011 and tried to look at what the potential impact of the percentage rate of hemolyzed specimens.

The target rate versus the 2 percent versus the Cleveland Clinic rate which is about 12 percent and versus that mean incidence of what's reported in the literature, anywhere from 6 to 18.

And based on those studies each percentage of hemolysis probably accounts for 300 redraw of labs, 300,000 of labs.

There's wide practice variation. Who draws the blood, how they draw the blood, the equipment utilized, and really currently no standardized approach or roadmap.

Because of this issue the Emergency Nurses Association also developed a clinical practice guideline on the topic called Prevention of Blood Specimen Hemolysis in Peripherally Collected Venous Specimens. And that's also included in our packet that you have there.

So, based on our studies. Like when we first studied this the percent of moderately

hemolyzed in our department was 18 percent.

And through a number of process improvement projects we were able to finally reduce that to below the 2 percent.

The poor quality specimens is the whole cause. There's a lot of argument about -- and this is what this always comes down to is the ED will complain that it's the lab's problem, and the lab says no, it's the ED's problem.

And on this count the lab is probably correct because the poor quality specimens are mostly due to pre-analytical errors, meaning how the lab is drawn.

And when you have a lab that's drawn poorly it results in delays in initiation of care. And after it results in more delays, in prolonged ED stays and wait times.

There's the potential for incorrect and missed diagnoses, and of course there's an increased healthcare cost.

CO-CHAIR SEPTIMUS: That's an excellent review. Thank you very much.

Just to remind the committee this is a new measure and it's for trial use approval just so we go there.

Lisa, did you have a comment before we start with the evidence? Yes, Lisa.

CO-CHAIR THRAEN: And it's eMeasure.

CO-CHAIR SEPTIMUS: Yes. Wait a minute, is it an eMeasure? Okay, I stand corrected.

MEMBER MCGIFFERT: Can you -- I'm sorry if I'm asking you to repeat yourself.

DR. PHELAN: That's okay.

MEMBER MCGIFFERT: Can you talk to me about the harm to patients when this occurs? I mean, I understand that they have to have a retest. But what are the other harms? And maybe missed diagnosis?

DR. PHELAN: The possibility that you could be misdiagnosed with a hemolyzed specimen that's reported out.

The potential at least from my perspective from the ED when I see someone who comes

into the emergency department and the initial lab reports to me as a potassium elevated 6 or 6.5 I immediately have to get an EKG. I immediately start the process until I get the repeat draw to come back because I'm still worried that is it hemolyzed and causing just a little bit of bump in potassium, or is it hemolyzed and causing a significant bump in potassium.

And I start an initiation of care for elevated potassium. Patient gets put on a monitor. I start giving insulin and glucose which can have repercussions. The insulin could drop the blood sugar.

I start giving other medications like Kayexalate which can cause some significant diarrhea. That's the whole purpose of giving it.

And if I get the repeat test back and it's normal I've done a whole bunch of stuff that I probably didn't need to do.

But, because it's such a potentially life-threatening problem, hyperkalemia, we jump on that right away.

The main thing that I see as the potential harm to patients is the delay in care. And with our data set we were able to identify that there's about a 50-minute delay in care for discharged patients, and about a 23 or 24-minute delay in disposition for admitted patients.

Based on the fact that you have to do a redraw. And when you do a redraw that means that the nurse that was actually taking care of someone else is pulled off that patient to come and do the redraw on this other patient.

And we did a time survey about how much that costs. It's anywhere from 10 to 12 minutes of a nurse's or a medic's time pulled off of other duties they could be doing to redraw a specimen.

MEMBER LAWLESS: This is Steve Lawless. Isn't this an American College of Pathology Q-probe?

DR. PHELAN: Yes.

MEMBER LAWLESS: And are they involved? I mean, in terms of the development. Did you solicit opinions from them in terms of

making sure this reconciles?

DR. PHELAN: We definitely have opinions from them. Based on some of our preliminary discussions, Dr. Howanitz, and I may have cited him in here, published two reports based on hemolysis. And if I didn't cite those I can send those to you.

But based on some of our preliminary discussions early on when we were looking at this topic Dr. Howanitz went back and sent a Q-probe just to identify the prevalence of the problem in laboratory medicine.

And those were published pretty recently, like in 2015. I'm not sure if I cited them in the citations, but I'm almost certain I did because there were two publications based on that.

But we have been in communication with them about this topic right through because of our cooperation with our lab medicine.

Because American College of Pathology runs the Q-probe and the CAP program. So they're

fully aware of this.

And they support it on the phone calls and things like that.

Unfortunately lab medicine and the CDC were a little hesitant to jump onboard to being part of this measure process because I don't think they were really informed enough about what it entails and what it involves.

I tried to pull them in as co-authors. They were kind of more in the standoff mode and wanted to kind of see where it went.

But they understand the significance of the problem, particularly lab medicine. They see it as one of their highest priorities.

CO-CHAIR THRAEN: I have a quick question for clarification purposes.

So, the measure that you're bringing forward is specific to the potassium samples.

DR. PHELAN: Correct.

CO-CHAIR THRAEN: When you quoted your 18 percent rate of hemolysis, was that specific only to potassium, or was that in general?

DR. PHELAN: That was specific to potassium. And that was what the data that we used to obtain a CDC-funded project grant or cooperative agreement.

So, when -- just to give you a little bit of the story, Gary Procop, one of our leaders in lab medicine, was approached by CDC for this, but he knew the issues and he knew the issues were mostly in the ED.

He provided that data for me and I was kind of shocked because I kind of thought it was a problem, but I didn't realize for at least our ED it was that significant.

MEMBER WEBB: Hi, this is Kendall Webb. So, I just want to clarify for the committee here. So, on every patient that you get a hemolyzed K back on you start all of that treatment? Or is that another test that you can do to not have to start that treatment until you get the final K back?

DR. PHELAN: So, if the lab sample is hemolyzed, and at our institution it's a little bit different. Many of the labs report this out

differently.

Our lab, and I wish my colleague Ed were here, but he wasn't available to attend, our lab gives us two results based on the hemolysis index.

If the hemolysis index is I believe greater than 300, and that's just a reading of their thing, they will not report the result. They say you have to redraw it because the variability in results is so off that they can't guarantee a 1 on the result.

So, between 80 and 300 they report it as the potassium sample is hemolyzed. Please use the result with caution because we can't guarantee where this result actually is. But they will report that out.

So, if a sample comes to me and it's hemolyzed and the result is normal I don't worry so much although the potential is there that that result could be significantly lower and the patient could have hypokalemia that I'm unaware of.

If the result comes back to me greater than 5.5 which is I think in our lab 5.5 is

considered hyperkalemic and it's a dangerous lab result I immediately begin treatment.

That's not every single hemolyzed specimen, it's just the ones that let's say the patient's true potassium was 4.5 and I get a hemolyzed specimen and that's going to kick it up another point. And so now it's reading it as 5.5.

Because hyperkalemia is such a life-threatening problem I would begin treatment on that even before I got a result back.

Now, in our ED we have the potential to get something called the point of care potassium. And my inclination is the reason we got the point of care lab was because of our high potassium result.

The point of care lab does not report out whether there is hemolysis or not. It is done on whole blood and there's no way on our machine to give us a thing that says hemolysis or not.

So I could redo that test immediately within five minutes and get a quick result back, but I have no idea and I did not know this until

I got involved in the project. I have no idea if the accuracy of that number is correct due to the fact that it doesn't report out if the sample was still hemolyzed or not.

So, there is still a potential danger that I could be either under-treating a patient or over-treating a patient based on whatever that result I get from the point of care lab.

Most times if the potassium result comes back still high, 5.5 again even on the point of care lab, the process that I used, I would send off another lab sample to get a more accurate result from the lab to be sure that I need to be treating hyperkalemia.

But I don't delay the care, I just begin treating hyperkalemia because it is such a dangerous, potentially life-threatening illness.

MEMBER WEBB: So is there a non-lab test that you could use to help you understand whether the hyperkalemia was causing danger in a patient?

DR. PHELAN: Not really, not that I'm

aware of. We get EKGs to initially see if there's EKG changes. Because there's subtle EKG changes that start with Q waves and end up with sinus wave which is a death wave basically because they have no cardiac function once the potassium reaches a certain level.

So, the non-laboratory test that we get to kind of confirm, but hyperkalemia does not correlate 1 to 1 with elevated T waves, or hyperelevated T waves. So it's a poor man's test to see if your hyperkalemia is really hyperkalemia and causing cardiac malfunction.

CO-CHAIR SEPTIMUS: Anything else, Kendall? Kendall's our Gator here, so.

Anyway, I think we've heard a lot of the evidence and so I think we're ready to vote. And they have provided enough information and review that we could rate this as high, moderate, low, or insufficient.

So, I'll turn it over to the voting wizard.

MS. QUINNONEZ: Yes. We are now

voting on measure 2983 Potassium Sample Hemolysis in the Emergency Department. Voting is now open for evidence.

Option number 1, high. Option number 2, moderate. Option number 3, low. And option number 4, insufficient.

All votes are in and voting is now closed. For the evidence of measure 2983 -- 6 voted high, 11 voted moderate, 1 low, and 2 insufficient.

CO-CHAIR SEPTIMUS: We'll move onto gap, but Kendall I think has her flag up. You have your hand up.

MEMBER WEBB: So, I was just a little bit confused, and this is really for the committee, not for the measure itself.

When they say it's a trial measure?

MR. LYZENGA: So this is something that Jason talked a little bit about earlier. So he's presented some evidence here, but he was not able to get the sites to do the like reliability and validity testing, that kind of thing. And not a

lot of information I think on gap as well.

So what we would do is if we want to approve this for trial approval it would not have endorsement status, but would have trial approval status.

He would then have to bring it back within three years with the testing results.

CO-CHAIR SEPTIMUS: Yes, he'd have to go through the full endorsement process. That's correct.

So, the next thing is --

MEMBER MCGIFFERT: So, I'm a little confused. So, if we felt that this was a measure that we wanted to go through the trial use then we would vote against it being endorsed. Or we would vote it low to be endorsed. No.

CO-CHAIR SEPTIMUS: We wouldn't vote for endorsement.

MEMBER MCGIFFERT: We wouldn't vote for endorsement, but if we voted low let's say on reliability and validity and gap and all that what does that do?

MR. LYZENGA: Well, we wouldn't actually vote on reliability and validity.

MEMBER MCGIFFERT: We won't vote on it.

MR. LYZENGA: Right.

MEMBER MCGIFFERT: Got it.

CO-CHAIR SEPTIMUS: Thanks, Lisa. Okay, gap. Is there anything else, Iona, that you wanted to say on gap that hasn't already come out in the conversation?

CO-CHAIR THRAEN: Just that the data that they submitted for gap is specific to their institution, Cleveland Clinic. And the graph before you is their analysis of where they started and then how they improved over the course of time indicating a gap.

Then also they talked about the literature in terms of the differences of hemolysis. And it varies across institutions. So I thought that they actually did support documenting a gap in performance.

CO-CHAIR SEPTIMUS: Comments. Yes, Yanling.

MEMBER YU: I'm confused with this figure. In 2013 the hemolysis rate is about 13 percent and it went down to 2 percent rate in 2015. That's an improvement, right?

DR. PHELAN: Correct.

MEMBER YU: So how did that happen? First question. And then does it still show a gap at all? When we don't have this measure it shows improvement.

CO-CHAIR THRAEN: So the developer, do you want to respond? You used the measure, correct? It was my understanding that you used the measure in your institution to see where you stood, and then you worked on an improvement process that you said there were multiple improvement efforts to achieve the goal of 2 percent which is the pathology's national standard recommendation.

DR. PHELAN: Correct. And this is in our institution. I'll give you a little bit of a briefer on what our ED is.

We have a combined residency with Metro Health Medical Center, the level 1 trauma center

across town, and the Cleveland Clinic main campus. So we share our residents.

I brought one of the residents onboard because he said, "I cannot stand working at Metro because one-third of my lab samples are hemolyzed and I have to wait for another sample to be done." So we're in the process of fixing their hospital across town.

Same issue with UH. A colleague of mine was like hey, this is killing us. And I'm like well, I have the answer for you if you'd like to try it.

So, at our institution, just the main campus, we started off at that high 15-16 percent rate. We collected data over the course. We did a couple of process improvement projects one of which was, as you can see I think it was in February, significantly dropped hemolysis.

We had to change some things at our lab and perform a couple of more process improvement projects to see if they had an impact. They had very little impact, but the replacement that we did

with our equipment significantly reduced our hemolysis down to 2 percent.

That is not to say that there's not a significant problem amongst our other hospital institutions. Within our own health system there is a significant gap there because I know that the rates are anywhere from 8 to 4 percent across our health system.

Four percent in our freestanding EDs. The average is about 8 percent in our standard hospital attached EDs. That's just in our health system alone.

I suspect across the nation there's great variability in this. And from presenting this at different scientific meetings people have come up to me and said we had a terrible problem and we did this. And oh, we fixed it doing this. And we did the same thing you did.

So there are hospital systems out there that have recognized this as a problem and tried to address it, but there's no uniform standard across the country about measuring it, and

collecting it, and being held accountable for it.

MEMBER YU: Okay, so I'm assuming the rates go up temporarily during the April and June period, it's because you were doing some other tests to try to refine your algorithm or whatever, your techniques, right?

DR. PHELAN: Correct. I'll tell you what happened. We looked at -- based on our nursing recommendation we looked at replacing our large volume 600ml collection tubes with 2ml collection tubes. We did that for one week.

We had some problems in the lab with quantity, the lab labels were covering the tube completely so the lab couldn't see inside.

So after we did it for a week we had to actually pull the small tubes back and replace them with our large tubes which you can see from that data it immediately bumped back up into the, I don't know what range, 14 for the combined hemolysis rate.

And come August when the lab was ready to switch back over we switched back to the small

tubes and it dropped the rate significantly back down to 2 percent.

And it's continued since. I mean, I continue to collect data on it which is ending soon though. But it's still at about 2 percent.

MEMBER LAWLESS: To the developer, I may have missed this in your literature. Is there literature in terms of the gap that supports or shows interventions that were done -- potassium comes back hemolyzed. Intervention was done, however, and then repeat was done, and they said oops, we shouldn't have done that.

Is there anything in terms of timeliness of delaying care as a result of this, or interventions done inappropriately because of this?

DR. PHELAN: No, not that I could find. But if you ask any ED physician who's done this they've had it come back where they're like oh, great, it's 4. It wasn't high, it was hemolyzed.

But I don't think anyone's done a survey or an analysis of it. And it becomes very

difficult to do sometimes because it would be hand done.

MEMBER WEBB: So, are you guys straight sticking every patient that comes in the ED now since that's the number one thing that supposedly helps?

DR. PHELAN: No, but there are EDs that have done that. And there are publications that have done it and it drastically reduces it.

I know about a hospital in Sarasota, Florida, and let me give you a little background on this.

ED nurses and medics are allowed to put an IV in. Phlebotomists are not. Some hospitals have gone to we're going to hire phlebotomists. It's very costly.

So, this hospital I know of in Florida started that way, but then they said we can't keep doing this even though their hemolysis rates dropped to zero because all a phlebotomist can do is straight stick a patient. They're not allowed to draw blood from an IV.

That hospital trained their nurses that they want all labs done via straight stick. Our nurses and a lot of people in the emergency nursing world have grave concerns with that because that requires, and if you've ever had your blood drawn and then had to get poked again for it, you'll know nobody likes that. From the nursing perspective they hate it.

So, our nurses, when we approached this project my plan was to just switch everybody over to straight stick and do that.

There was great consternation from our nurses. We have a large tertiary care population that is very difficult to stick let alone poke once to draw labs, two to get an IV. So there was pushback from my nurses.

And I initially was very much, like I tried to work the best I could with leadership to say can we do this. There was great pushback.

There was a paper published by a person named Dietrich out in Wyoming and I think I included it in the publications.

And there's arguments in this literature all the time, one poke or two pokes. The lab medicine people want a perfect lab specimen. They want two pokes. The nurses that have to sit there with a crying, hurt patient have to poke them twice do not want to poke these patients twice.

So, this Dietrich publication said one poke or two. One poke works just fine. I'm not sure I know what the issue is.

When I contacted him because I could not figure out how he had such a low potassium rate for blood draws through an IV because it's typically much higher.

He didn't know, and then we started exploring the small tube. And I re-contacted him and he goes yes, we've been using 2ml tubes for about three years. And I said that's probably where you're getting your low hemolysis from.

And we published a letter to the editor in Journal of American Nursing that I can forward to the committee if they want to see it describing

that exact issue.

But there is a big dichotomy between the lab people and the ED nursing and ED clinicians about one poke or two. And it continues to be kind of fought out in the literature and sometimes kind of with sparks flying.

CO-CHAIR SEPTIMUS: So that's an interesting topic. And just to fill in some of the gaps.

So, we have HCAHPS scores which everybody's concerned about because it fits into value-based purchasing.

We also have the issue, by the way, of blood culture contamination rates which is directly related to what you stated. So this is a very interesting topic. But the topic in here, we're talking about hemolysis.

So I think we're ready to vote on the gap. So let's do the gap.

MS. QUINNONEZ: Voting is now open for performance gap of measure 2983.

Option 1, high. Option 2, moderate.

Option 3, low. And option 4, insufficient.

All votes are in. Voting is now closed. For the performance gap of measure 2983 30 percent voted high, 60 percent voted moderate, 10 percent voted low and zero percent insufficient.

CO-CHAIR SEPTIMUS: Okay, so we're going to go to reliability. There's no testing here so what we're going to be discussing is their numerators and denominators as a measure only. Does that make sense to everybody?

MR. LYZENGA: I was a little wrong when I said that to you, Lisa, earlier. We will vote on reliability, but only on the specifications part of it.

CO-CHAIR SEPTIMUS: The specs.

MR. LYZENGA: The precision.

CO-CHAIR SEPTIMUS: There's no testing so we can't -- okay.

CO-CHAIR THRAEN: So, I may be confusing two items. So is this the measure in which trying to capture the information in an electronic health record wasn't working very well

and you needed to go to the lab information system to capture the data? Can you clarify that for me from the developer?

DR. PHELAN: Sure. No, we were able to capture it from Epic and Sunquest. Both are ONC certified at our institution.

We did not buy the Epic lab information system. So we had to have a dual ONC certified. And we got it both from Epic and from Sunquest which is our lab information system.

CO-CHAIR THRAEN: Okay, thank you. The question is did they correlate. Did the two sources correlate with one another.

DR. PHELAN: Not exactly. And our plan was when we do the reliability testing to further identify why the Epic pull didn't quite match the Sunquest pull.

And when we were in the midst of the research project we were doing it, but when the research project ended, you know, when funding dries up there's no more anyone willing to do any more work on it.

But we were identifying, we were going through month by month to try to see why Epic included some cases and Sunquest didn't. And I actually have a meeting this afternoon to talk to our business intelligence people and Sunquest, I finally got them in the room together, to look at why we're potentially not having an exact correlation.

But if I remember correctly the numbers were off by 1 or 2 percent. So the gross hemolysis for 2014 was 4 percent in Epic and 3 percent in Sunquest.

And I don't know if it was the data asks were a little bit different. So we're trying to ask if they can get the same data ask. You know, we want just main campus ED, just the month of January, all potassiums minus point of care potassiums. Because point of care potassiums don't show up on the main lab data system, on the Sunquest data system.

CO-CHAIR SEPTIMUS: Okay, just to keep everybody -- the denominator here is all patients

who come to the ED and who the potassium is ordered. And the numerator includes patients in whom a lab potassium sample is reported as hemolyzed, correct?

DR. PHELAN: Correct.

CO-CHAIR SEPTIMUS: Okay, I just wanted to make sure everyone understood what the eMeasure is. Okay, Yanling and then Steve.

MEMBER YU: Yes, thank you. I guess I'm confused about the denominator and how it defines uses.

I'm looking at the article that you cited about -- in a previous section of the proposal.

This article cited the effectiveness of practice to reduce blood sample hemolysis in EDs, a laboratory medicine best practice systematic review and meta analysis that's recorded in the developer's documentation.

But the conclusion is to use a new straight needle venipuncture instead of IV to start is an effective way to reduce hemolysis rates in

the ED.

That seems like this documentation shows nothing about whether it's a potassium sample or not. So, maybe I'm just confused. It seems like it's how you draw the blood rather than whether it involved just potassium sample. Am I wrong on that?

DR. PHELAN: No, no, you're right on. It's the technique of the blood draw that contributes to the hemolysis.

And the laboratory medicine's best practices looked at hemolysis in general. There are 39 different lab tests that are affected by hemolysis, many of which I don't care so much about. The one I care most about is the potassium.

It's a frequently used test and we often get the report out that it's hemolyzed some way. And I have to alter my practice or delay the care to my patient based on that result.

So it does affect other lab tests like bilirubin, type and screen, coagulation studies. I specifically focused on ED hemolysis because I

wanted a pretty simple measure to capture and results can be captured easily.

CO-CHAIR SEPTIMUS: So another analogy here, and Steve is next. One of the things that most organizations measure is blood culture contamination rates.

And one of the highest areas of blood culture contamination rate happens to be the ED. And it directly relates to how blood is drawn.

And this is just an analogy. So people track that. And when they see a high blood culture contamination rate they have an intervention.

And I think that's sort of what we're hearing here. We have a high hemolysis rate. There are best practices for how blood is drawn, and maybe what size tube it's drawn in. And that hospitals can track that. And it does have an impact on the validity of a number of laboratory tests.

I'm just trying to give you some parallels here. So Steve and then Albert.

MEMBER LAWLESS: Going back to your

specifications about your numerator, two questions.

One, are you talking about the Epic lab system, LIS system versus the Sunquest LIS system? Or Epic pulling from Sunquest? Because it's just a query pull.

And the second question is about the numerator. There's hemolysis and then there's hemolysis. Both the laboratory person who's looking at the specimen versus the machine in the lab may say hemolysis. How do you specify or quantify the degree of hemolysis to be clear on this?

DR. PHELAN: So, your first question was again?

MEMBER LAWLESS: There's an Epic laboratory information system and there's also the Sunquest lab information system.

Is your differences between that, that you're testing one LIS system versus another? Or are you saying Epic's pull from Sunquest is showing a difference?

DR. PHELAN: I think it's the latter. Our institution did not acquire Oliver or whatever the lab information system from Epic was.

So, data is pulled or pushed into Epic from Sunquest, and it populates a field. I can go to my business intelligence which I have to admit is quite an oxymoron at least where I'm at, and say can you pull me data on these patients. And it's a struggle to get that data. That was one of my largest things.

Lab medicine, the week I got the grant, through their software system and their -- through their Alto software, I just specified every ED patient that gets a lab draw test from this period to this period, and I want it on a monthly basis. They gave me that almost immediately.

And it correlated -- the first time we got it we hand went through it. We're like yes, there's about that many. We were off by one patient maybe and we never could find out why we didn't have that one patient.

But when you were looking at twenty or

thirty thousand lab specimens it didn't matter. I was happy that the Sunquest lab information system gave me data right off the bat.

Epic is data that Sunquest pulls in through some middleware, and I'm not sure exactly how, it populates in Epic and then I have to get a SQL programmer to say, well, what do you want. Okay, I'll pull all the lab tests for you. And then they pulled some point of care lab testing that I didn't want. And so we went with multiple iterations.

It was not the lab information system that you can purchase with Epic. So it was just standalone Epic versus Sunquest lab information system.

Our Epic as a medical record is ONC certified and our lab information system Sunquest is ONC certified. So we can pull from either one. I would start with our lab information system and use that as the gold standard.

CO-CHAIR THRAEN: So I just want to refocus you on this question about how you decide

that something has been hemolyzed. There's like three different ways to say this term. Because it's qualitative in nature.

DR. PHELAN: At our lab they provide you with a hemolysis index. And based on that hemolysis index if it's between 30 and 80 they just give out a report, there's no hemolysis.

Between 80 and 300 they call it moderately hemolyzed. They give you a result, but they also put a comment in the comment field that says hey, be cautious in interpreting this. It's not 100 percent that it's right on the money, but it's there.

If the lab sample has a hemolysis index of greater or equal to 300 there will be no result in the result box and it will say grossly hemolyzed. Please resend the sample.

MEMBER LAWLESS: So which one do you use for this numerator?

DR. PHELAN: What's that?

MEMBER LAWLESS: Which of those three grades do you use in the numerator?

DR. PHELAN: I use a combined HK which is moderately hemolyzed and GK which is grossly hemolyzed. I use the total hemolysis as the numerator. I don't separate out gross and moderate.

Because anytime you're having hemolysis it potentially could affect your result.

CO-CHAIR SEPTIMUS: Okay, so here's where we want to say are the data elements clearly defined.

And I think one of the questions is not the denominator. The question is are all ED laboratories able to give us percent -- well, based on your definition of moderate to severe, is that something that's commonly reported and commonly done.

DR. PHELAN: No, hence we did a total hemolysis. So all they would have to do is provide you with whatever hemolysis rate they're getting.

CO-CHAIR SEPTIMUS: So, I guess during this trial period as you test this more we might get a better sense as to how good that numerator

is and the definition of the numerator? Because I think that's the soft point here.

It's not the concept and it's not the denominator, it is whether or not this measure is reliable across other multiple systems to give us reliable data that's actionable so that this can be tracked and action can be taken to reduce hemolysis. Am I expressing that okay?

DR. PHELAN: Perfect.

CO-CHAIR SEPTIMUS: So why don't we go ahead and let's vote on the specifications in terms of reliability.

MS. QUINNONEZ: Voting is now open for measure 2983 for the measure specifications for eMeasure approval for trial use.

Option number 1 is high. Option number 2 is moderate. Option number 3 is low. And option number 4, insufficient.

DR. PHELAN: And can I say one thing before you continue? ED hemolysis and hemolysis in general is monitored by every single lab system across the country. They all know about it. They

all know it's a problem.

The exact definitions may be off because everyone has a different machine, or different things, but they all measure hemolysis and they all know ED is their number one source of hemolyzed lab specimens in their system.

CO-CHAIR SEPTIMUS: Appreciate that clarification. Obviously we'd like to have the measurement be consistent across most EDs.

But thanks for that clarification because you're right. It's like blood culture contamination, they all monitor it. Blood culture contamination is a little more straightforward I think than how they monitor hemolysis.

DR. PHELAN: I agree.

CO-CHAIR SEPTIMUS: But I understand what you're saying. So let's go ahead and let's start over again.

MS. QUINNONEZ: Okay, voting is now open for measure 2983 on the measure specifications for eMeasure approval for trial use.

Option 1, high. Option 2, moderate.

Option 3, low. And option 4, insufficient.

All votes are in. Voting is now closed.

For the measure specifications of 2983 16 percent voted high, 84 percent voted moderate, zero percent for low and zero percent insufficient.

CO-CHAIR SEPTIMUS: Okay, so we're obviously not going to do validity, but we are going to talk about reliability and usability -- feasibility. Whatever, you know what I'm talking about. Feasibility and usability. So, Iona.

CO-CHAIR THRAEN: So, feasibility, as you've already heard most or all lab systems have the capacity.

This is an eMeasure that can be easily pulled. Not without it sounds like some cleaning and some clarification that needs to happen depending on what system you're talking about.

But if you're dealing directly with your lab systems it sounds like that's an easier path than your EHR. So I think it's highly feasible.

CO-CHAIR SEPTIMUS: Seeing no comments let's go vote.

MS. QUINNONEZ: Voting is now open for the feasibility of measure 2983.

Option 1, high. Option 2, moderate. Option 3, low. And option 4, insufficient.

All votes are in. Voting is now closed.

For the feasibility of measure 2983 61 percent voted high, 39 percent voted moderate, zero percent for low and zero percent insufficient.

CO-CHAIR SEPTIMUS: Michelle, are you still on the line?

MEMBER SCHREIBER: Yes, sir.

CO-CHAIR SEPTIMUS: Okay, we don't want to forget you. So if you need to raise your hand can you email Drew so we can make sure that you're recognized, okay?

MEMBER SCHREIBER: Yes, thank you.

CO-CHAIR SEPTIMUS: I should have asked you that before, but thanks.

So the next one is going to be

usability.

CO-CHAIR THRAEN: So it's not currently being used publicly or for accountability, but they do plan to recommend this for accountability in the future.

They also indicated that their planned use includes public reporting, public health disease surveillance, quality improvement with benchmarking and quality improvement in internal benchmarking across organizations. So they have very strong hopes to use this in the future for many uses.

MEMBER ARDIZZONE: Can I ask a question? Just a practical question to the developer. I see unintended consequences. Is there any price difference for a 2cc vial versus a 6cc vial? Or any sort of those things that maybe small EDs can't, you know, manage, or workflow changes?

DR. PHELAN: Not that I'm aware of. I think they're equivalent. If you'd like I can actually reach out to our nurse director to see if

there's a cost difference between a 6ml tube and a 2ml tube.

My impression is because it's smaller it should probably cost less. Less glass, less plastic.

MEMBER ARDIZZONE: Not necessarily.

DR. PHELAN: I know, I know.

MEMBER ARDIZZONE: Does it change the measure? I was really just interested if there's some sort of performance change out of this that should be adopted nationwide can every institution conceivably apply it.

DR. PHELAN: I have a feeling after going through what I went through that most EDs are going to opt for the least amount of resistance which is replacing their 6ml tubes with 2ml tubes because it works great.

And there's very little up front communication that has to be done other than with labeling process and things like that.

If you're changing to a straight stick process there may be an increase in cost there.

And I'm not sure how, but because you're doing a two-stick process. You're doing one stick with a straight stick needle and drawing the labs out. Then you're adding more equipment by putting an IV in.

So the less equipment used by doing IV, there may be less cost associated with it, even if there's a potential increase in the cost of a smaller tube.

But if the committee would like I can reach out to our nurse director and our purchase person and find out if there's a cost differential between the large and the small tubes.

CO-CHAIR SEPTIMUS: So, these comments about usability. So we're supposed to look at --

MEMBER ARDIZZONE: I thought that was considered a potential harm or unintended consequences.

CO-CHAIR SEPTIMUS: Well, that was the other thing, potential harm or unintended consequences.

MEMBER ARDIZZONE: Yes.

CO-CHAIR SEPTIMUS: So, Charlotte and then Kendall and then Yanling. Charlotte.

MEMBER ALEXANDER: Just for my information because I don't know oftentimes when you're drawing lab works you're getting more than just the potassium.

So is the 2cc tube a sufficient sample that you can get the other tests you need to get? Or do you have to get another tube to get enough blood to get the other tests?

CO-CHAIR SEPTIMUS: Developer, did you hear that question?

DR. PHELAN: Yes. We draw three 2ml tubes where we used to draw three 6ml tubes. And we've had no problems with the sufficiency of the material.

Now, I can't speak to other lab systems and other machines, but from our perspective I haven't had a problem.

It was an argument whether we needed one more tube because of a request for more tests. Like there's an add-on that happens in the

emergency department. You're working someone up and you want to add something on.

We looked at it for about a month and there was not an increased request for add-ons during that time period at our lab.

We went from considering four 2ml tubes back down to three, and we've just kept at three since, and I haven't heard a problem since.

MEMBER WEBB: So, I just want to actually answer from my knowledge base where the extra cost would be is I can tell you if we went to a 2ml system we would need to change all of our lab label printers, and all of the interfaces that go to those lab label printers because our lab labels are for a 6ml tube. So it's not just the cost of the tube, it's the cost of the EMR at this point, and the equipment that goes with the EMR potentially. So that would just be something sort of outside the box.

CO-CHAIR SEPTIMUS: Yanling.

MEMBER YU: Thank you. The question, two short questions for the developer.

First of all, the planned use includes public reporting. So my first question is what do you have in mind about the public reporting. Any future plans? Just share your ideas.

And a second question is for quality improvement as you explained that this could cause harm and could have bad patient outcome. So do you have anything in your mind that could -- down the road to tie to some outcome measure, you know, go from this eMeasure. Thank you.

DR. PHELAN: The first thing on the public reporting. Because it's such a ubiquitous problem I could actually see it as being a nursing, an ED, or a hospital measure.

And because it significantly affects patient throughput and I think ED throughput, it's still one of the core measures that they're looking at in either IPPS or OPPS. I think it's the Outpatient Perspective Payment System.

I definitely see it as impacting things like patient throughput. Because when you start duplicating work and redoing it, and from an

efficiency perspective with more and more focus on the cost of what it does, having the cost and the labor cost distributed and needed to move around, I see this having a significant impact on throughput and efficiency all around.

CO-CHAIR SEPTIMUS: Okay, I think we're ready to vote on usability.

MS. QUINNONEZ: Voting is now open for the usability and use of measure 2983.

Option 1, high. Option 2, moderate. Option 3, low. And option 4, insufficient information.

All votes are in and voting is now closed. For usability and use of measure 2983 24 percent voted high, 76 percent voted moderate, zero percent for low and zero percent for insufficient information.

CO-CHAIR SEPTIMUS: Okay, so I think that's the last question for this because we're not voting on endorsement. So we want to thank the developer for a -- oh. Was there one more question? Oh, I'm sorry.

Okay, approved for testing. Trial use.

MS. QUINNONEZ: Voting is now open for overall suitability for eMeasure approval for trial use of measure 2983.

Option number 1 is yes. Option number 2 is no.

All votes are in. Voting is now closed. For the overall suitability for eMeasure approval for trial use for measure 2983 100 percent voted yes.

CO-CHAIR SEPTIMUS: Okay, so Pat has a comment, then we're going to go to public comments, and then we're going to go to lunch early. And I'll tell you when we're going to restart. Everybody with it? Pat.

MEMBER QUIGLEY: Thank you, and I will be brief. This is Pat Quigley's voice for the developer on the call.

I just want to applaud you and thank you for all that you did to teach us as you went through this.

I just hear your joy of improvement and your joy of interdisciplinary involvement and working with the nursing staff there as well as the lab staff. And I just want to thank you so much.

DR. PHELAN: Oh, you're welcome, that was nice of you to say. And it was really a team effort. I give most of the credit to Emory Kavach because she was our nurse director in the ED at the time.

And she just bought hook, line and sinker into this without a whole lot of effort on my part. I mean, I was really shocked because I was expecting a lot of pushback. But they also pushed back to doing the straight stick so I was a little upset at that, but it turned out great regardless. And it was a great team effort with lab, ED and nursing. It really was.

MEMBER QUIGLEY: Thank you.

CO-CHAIR SEPTIMUS: Lillee? Okay, public comment. Operator?

OPERATOR: Okay, at this time if you would like to make a comment please press * then

the number 1. At this time there are no public comments.

CO-CHAIR SEPTIMUS: Okay. So, we'll finish with Lillee and then we're going to go to lunch. And then let's try to come back at 12:30 and we'll try to make sure that we finish on time and get all of our business in. Lillee.

MEMBER GELINAS: We were mentioning at dinner last night, and thank you for that very much, that there are several changes to the biographies for this committee that are needed. Organizations, titles, phone numbers, the whole nine yards.

And I know we're going to start losing people soon. So I don't know what the protocol and process is for NQF, but I just wanted to make sure maybe over lunch we could get that done.

MS. QUINNONEZ: That would be awesome. Actually, if you could send your new biographies and titles to the patient safety email box we'll go ahead and update that for you. Thank you.

CO-CHAIR SEPTIMUS: Thanks for

bringing that to our attention. Okay, lunch is now served.

(Whereupon, the above-entitled matter went off the record at 11:59 p.m. and resumed at 12:30 p.m.)

CO-CHAIR SEPTIMUS: Our next measure is going to be 3025 Ambulatory Breast Procedure Surgical Site Infection Outcome Measure. And the CDC is the developer and they are on the line. And Dr. Alexander will be the discussant. And I'll turn this over to Iona for the first part to moderate.

So, the CDC developer, can you announce yourself, and tell us your name, and then go over your measure?

DR. POLLOCK: Yes, this is Daniel Pollock at CDC in Atlanta. We're very pleased to work with you today on the measure proposal we submitted.

It's a proposed measure that we co-developed with the Ambulatory Surgery Center Quality Collaboration and also the Colorado

Department of Public Health and Environment.

The measure itself provides a summary statistic. In accordance with other CDC stewarded measures we're using the standardized infection ratio to summarize the observed to predicted surgical site infections following breast surgeries in ambulatory surgery centers.

This is a risk-adjusted measure. It's an outcome measure. And this is a procedure with SSIs that was selected because it is -- breast procedures are the highest volume surgical procedure reported to CDC's National Healthcare Safety Network from ambulatory surgery centers.

And in the data that we have on the procedures that have been reported in from ambulatory surgery centers breast surgeries pose the highest risk of infection.

So it's a high-value target with prevention opportunities that has been developed in concert with the two organizations I've mentioned.

We currently are in the process of

building into NHSN a new outpatient procedure component. And the idea will be to have this particular measure serve as an initial measure of SSIs to be followed by other SSI measures that would be pertinent in the ambulatory surgery center arena. I'll stop there.

CO-CHAIR SEPTIMUS: Hi, Dan. This is Ed. I just want to say hi.

DR. POLLOCK: Hey Ed.

CO-CHAIR SEPTIMUS: Is this the first ambulatory measure that you've come up with?

DR. POLLOCK: This is the first ambulatory surgery center SSI measure that we have come up with, yes.

CO-CHAIR SEPTIMUS: That's what I thought. Thank you.

DR. POLLOCK: Yes.

CO-CHAIR THRAEN: Okay. We're switching. So Charlotte, you are the lead on this.

MEMBER ALEXANDER: So, the description was beautifully done by Mr. Pollock. And the level of analysis is at the facility.

The evidence is not overwhelming in volume but fairly consistent and is consistent with guidelines that we have for inpatient facilities as well, that the surveillance and reporting back does show a decrease in the instance of infections.

There is a CDC draft guideline for the prevention of surgical site infections which they have cited. And there are some articles also that they've cited.

The risk that's been identified has been listed as being as high as 30 percent which is a significant risk.

And as we know infections anywhere can significantly burden the system both financially as well as impact on the patients.

There was a five-year study of surgical site infections in the ambulatory surgical arena which showed a rate of about 2.8 per 100.

There's not been consistency in the rates that have been reported, but certainly they have been high enough to be of concern.

So I think that the data is there to

support the measure.

CO-CHAIR THRAEN: Any questions, comments? Let's vote.

CO-CHAIR SEPTIMUS: So we're voting on the evidence.

MS. QUINNONEZ: Voting is now open for the evidence of measure 3025.

Option number 1, yes. Option number 2, no.

CO-CHAIR SEPTIMUS: Just to remind you this is an outcome measure. That's why you're seeing this different.

MS. QUINNONEZ: Just to repeat, this is the vote for measure number 3025 Ambulatory Breast Procedure Surgical Site Infection, the SSI measure.

And option number 1 is yes and option number 2 for evidence is no.

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

The result for the evidence of measure 3025 is 100 percent voted yes.

CO-CHAIR THRAEN: All right, performance gaps.

MEMBER ALEXANDER: So, breast procedures compose almost 50 or 46 percent of the procedures done in ambulatory surgical centers.

And they comprise almost 55 percent of the reported infections.

The SSI risk is about 0.25 percent. So it's the highest volume procedure being performed and it has the highest risk of a procedure in an ambulatory surgical center.

This is particularly disturbing because as we look at the trend in healthcare we're moving more and more toward ambulatory procedures and away from inpatient procedures. So, I think this is a growing population that we really need to address.

They did stratify by age and gender, and showed disparities there. So I think there is an opportunity for improvement that's well demonstrated.

CO-CHAIR THRAEN: Questions.

MEMBER WU: Could you just clarify, is that 0.25 percent or 25 percent?

MEMBER ALEXANDER: 0.25.

MEMBER WU: One in four hundred cases there's a surgical site infection?

MEMBER ALEXANDER: Correct. So, ambulatory surgical infection rates are about half of the inpatient surgical infection rates. So inpatient is about 4 percent and ambulatory is about 2 percent.

And if you look at then the risk of a breast patient getting an infection that's what the 0.25 is.

MEMBER WU: Up here it says -- it's 2.5 or 0.25?

MEMBER ALEXANDER: It says 0.25.

MEMBER WU: Do we know what the actual numbers are? I was just confused further. Is it 4 percent and 2 percent, or is it 0.2 percent and 0.4 percent?

MEMBER YU: No, 78 divided by 30,787 you get exactly 0.25 percent.

CO-CHAIR SEPTIMUS: Dan, do you hear the question?

MEMBER WU: What's the general rate of infection for breast procedures, for example, in maybe inpatient, and then is that much higher in ambulatory, and is that higher than for other procedures?

DR. POLLOCK: I don't have in front of me the data on breast procedures among inpatients, but in terms of the relative risk compared to other types of procedures in ambulatory surgery centers among the procedures that have been reported into NHSN breast procedures have the highest risk of surgical site infection.

So they're a high-volume and relatively speaking high-risk procedure for an SSI in the ambulatory surgery center data that we have in NHSN.

And we have over 30,000 breast procedures reported in for the study period with 78 infections detected, reported in. So it's about 0.25 percent is the risk.

CO-CHAIR SEPTIMUS: Dan, this is Ed. So, I think if I remember the specs on this you're not just confined to deep organ space. This is also superficial, is that correct?

DR. POLLOCK: No, this is deep organ space. I'm sorry, it is superficial and deep organ space.

CO-CHAIR SEPTIMUS: So that's a little different. And the reason I ask you is that could the rate of 0.25 if you consider that we have a hard time capturing the superficial could the rate actually be higher?

DR. POLLOCK: Well, there are many reasons why it could be higher. These are the infections that are reported in.

We all know that some of the most challenging parts of surgical site infection surveillance are capturing the infections in the outpatient phase of care.

And by definition ambulatory surgery is done same-day surgery. So there is always going to be an outpatient surveillance challenge in

ambulatory surgery centers.

So, our conjecture is that the 0.25 percent is probably a low estimate of the true extent of the infection burdens being placed.

CO-CHAIR THRAEN: Steve and then Yanling.

MEMBER LAWLESS: Two questions. In your stuff you sent to us that rate was 0.25 but it went up to almost 28 per 1,000. So it's up to 2.8.

Is there a reference to that, or that variability? You're saying 0.25, but you also have said up to 2.8.

DR. POLLOCK: Well, there's literature as well as data that have been reported to NHSN. So we've done our best to summarize the literature as well as provide the actual surveillance data that we've received during the study time period.

MEMBER LAWLESS: Okay, thank you.

MEMBER YU: Maybe just a comment. I know in our state, in Washington the medical board is trying to draft up a policy that requires

ambulatory surgical infection be reported by physician.

Some of the cases came to us are exactly woman getting infected because of breast surgery in ambulatory centers.

The question we are wrestling at the time I remember was about what time window that physicians are required to report.

So I wonder do you have any thoughts on how you can define this superficial or deep infection based on the time window.

DR. POLLOCK: Yes, good question. So, we have specified that a 30-day time window for the superficial surgical site infections and a 90-day time window for the deep in organ space surgical site infections.

CO-CHAIR THRAEN: So, I just want to point out for those of you that may or may not know ambulatory surgical centers are the freestanding surgical centers. And they have not historically reported in national terms what their rates are.

They do report, often many of them do

report to their association. And they do have an accreditation body that they report their findings to.

And it varies by state. In Utah they are required to report sentinel events to us, but not quality measures.

And so this is sort of another wild, wild West world experience stepping into that environment and asking them to share their and become transparent on their outcomes as well as what we've done historically with hospitals and skilled nursing facilities.

Charlotte.

MEMBER ALEXANDER: So I have a question for the developer. I know that most of the time the reporting is done voluntarily by the physician. So a survey or a questionnaire is sent to the physician asking if he has infections.

It's not super common that the infections come back into the facility.

In the inpatient world if I have a surgery and it gets infected and they go to another

hospital there's a very good collaboration between hospitals as far as reporting a complication that's come in at a second hospital.

Is there a way that there is a communication between the inpatient world and the ambulatory surgery world where they can report infections that come in that are secondary to an ambulatory surgery procedure?

DR. POLLOCK: That's a great question. Certainly there are ways for that to happen and it does happen now.

There are ways to incentivize that type of communication. We're keenly interested in helping to incentivize that type of communication.

And much depends right now on the local practice and the network of infection prevention personnel in hospitals and their connections with the ambulatory surgery centers in their community. So it varies.

But over the long haul we definitely see opportunities to invigorate those types of communications and would value any type of

opportunity to ramp that up as soon as we can.

We think this is a step in that direction, that when we move into the quality measure arena on top of the six or seven states that require already SSIs to be reported from ASCs to NHSN.

We're on a trajectory where the visibility of surgical site infection in the ambulatory surgery center area is increasing and it's a problem as one of the previous comments alluded to that we know less about than we do and particularly with respect to SSIs in the inpatient arena.

And with the increasing volume of procedures done in the outpatient setting it's very important for us to bring under surveillance the high-volume high-risk procedures to begin with and move from there.

CO-CHAIR THRAEN: Shall we vote?

MS. QUINNONEZ: Voting is now open on performance gap for measure 3025.

Option number 1, high. Option number

2, moderate. Option number 3, low. And option number 4, insufficient.

All votes are in. Voting is now closed.

For the performance gap of measure 3025 37 percent voted high, 63 percent voted moderate, zero percent for low and zero percent for insufficient.

CO-CHAIR THRAEN: Charlotte, reliability.

MEMBER ALEXANDER: There was not reliability testing done on the measure or on the data.

There was reliability testing that was done on the risk stratification.

The model that they used was one where they focused on procedures from selected surgery centers in Colorado for the period of January to December.

They chose the surgery centers that had a minimum of 100 patients volume in breast procedures during that year's time.

They then looked for over- and under-reporting discrepancies and omissions.

There was no under-reporting that they found. There was one over-reporting that was a definition issue.

They had five facilities that entered total procedure duration for bilaterals rather than separating them out as two separate procedures.

They had a high percentage of the facilities where the procedure duration was incorrect. It was actually 95 percent where it was incorrect.

They had been using a protocol definition that had been in place prior to 2014. The measure year was 2014 and so it had just been changed over and the facilities had not done that change.

They felt it was highly reliable in identifying the SSIs because there was no under-reporting. But there was no testing.

MR. LYZENGA: We considered that to be

data element testing, reliability testing, or validity testing, the analysis that they did.

MEMBER ALEXANDER: Okay, good.

CO-CHAIR THRAEN: FYI, Colorado's ASCs are very proactive in this area so I'm not surprised that it was done in Colorado. They're much more organized in that state than any of the other states that I've seen over the years.

Any questions? Go ahead, Pat and then Laura.

MEMBER QUIGLEY: Thank you. This is Pat Quigley's voice for the developer. Thank you for those comments.

And I'm asking a question on behalf of Dr. Kimberly Applegate who was not able to be with us today. And she asked that we ask.

She's a radiologist, a pediatric radiologist and director of practice quality improvement in radiology at Emery University.

And her question in relationship to reliability is if -- in testing the reliability of the procedures is if the procedures were

image-guided procedures that were done.

DR. POLLOCK: Unfortunately we don't have that information. We do know simply that these were breast procedures in ambulatory surgery centers that met procedure code criteria. And we have ICD and CPT procedure codes.

And the time period that this was done the procedure codes did not -- were not captured at a level where even if they did provide the information about imaging we don't collect that.

MEMBER QUIGLEY: Thank you.

MEMBER ARDIZZONE: Can I just clarify because I overheard some of that conversation? She also wanted to make sure that the CPT codes and the ICD-10 codes that you captured would have MRI, like needle localizations, all those other breast procedures that are radiology assisted.

Because she read through your CPT and ICD-10 list and didn't see that those were captured.

DR. POLLOCK: My colleague, Kathy Bridson, just pointed out that needle aspirations

if this is the issue that we're talking about, if they don't have an incision they don't qualify for inclusion in the code list that we provide for facilities to use and that Colorado uses.

MS. BRIDSON: They don't meet our definition of an operative procedure if they don't have an incision.

CO-CHAIR THRAEN: Repeat that.

MS. BRIDSON: Our NHSN definition of an operative procedure is one that involves an incision, not a percutaneous or a needle --

MEMBER ARDIZZONE: Right. I think if I can clarify again -- again, I'm speaking for somebody else -- that it wasn't just -- that sometimes there's a combined procedure. So an MRI needle localization, and then you do a lumpectomy, or a sentinel lymph node biopsy afterwards.

So I guess what you're saying is if there's a procedure associated with the radiology component of it it would be captured in the CPT coding.

DR. POLLOCK: Yes. If there's an

operative procedure it would be captured in the CPT coding, but not the imaging per se.

MEMBER ARDIZZONE: Got it. But it would be captured if there was an incision afterwards.

DR. POLLOCK: Correct.

CO-CHAIR THRAEN: Other questions? All right, let's vote. Reliability.

MS. QUINNONEZ: Voting is now open for the reliability of measure 3025.

Option number 1, moderate. Option number 2, low. Option number 3, insufficient.

Option number 1, moderate. Option number 2, low. Option number 3, insufficient.

If you need to change your vote the clicker will capture the last number that you choose.

All votes are in. Voting is now closed.

For the reliability of measure 3025 60 percent voted moderate, 25 percent voted low, 15 percent voted insufficient.

CO-CHAIR SEPTIMUS: Dan, I have a question for you. It says you exclude very elderly patients.

DR. POLLOCK: Yes.

CO-CHAIR SEPTIMUS: And it says there at the top ages 18 to 108.

DR. POLLOCK: Correct.

CO-CHAIR SEPTIMUS: So, what's very elderly? We're having this discussion about the definition of elderly.

DR. POLLOCK: Elderly looks to be a higher and higher number every year for some of us. And so we're saying 109 is truly elderly.

CO-CHAIR SEPTIMUS: I'm safe for a couple of years.

CO-CHAIR THRAEN: All right, so validity.

MEMBER ALEXANDER: So this was a face validity assessment. There was a consensus process. There were 11 individuals with about 80 percent concurrence that the measure measures what it's intended to do.

And the majority agreed as well that it accurately reflects quality.

Risk factors were assessed and there was a univariate analysis and backward elimination.

It ended up using ASA classification and age as valid to adjust for the statistics. This was risk-adjusted using a statistical model.

They used the Hosmer-Lemeshow with a P equal to 0.66.

And with that zero shows no correlation. And the range is from zero to 1, so it's more than 50 percent.

The SEER was not calculated if the predicted value was less than 0.2. So of 138 facilities the SEER was able to be calculated for 70.

They looked at missing data on the ASA class. It was about 18 percent.

They looked at that population compared to the other population and felt that the crude risk in the missing procedures was not significantly

different from the risk of the others. It was 0.36 compared to 0.25.

And so I think this represents moderate validity.

CO-CHAIR THRAEN: All right. Pat, did you have a question?

MEMBER QUIGLEY: No, I'm so sorry. Thank you.

CO-CHAIR THRAEN: Any questions? Let's vote.

MS. QUINNONEZ: Voting is now open for the validity of measure 3025.

Option 1, moderate. Option 2, low. Option 3, insufficient.

Option 1, moderate. Option 2, low. Option 3, insufficient.

All votes are in and voting is now closed. For the validity of measure 3025 89 percent voted moderate, 5 percent voted low and 5 percent insufficient.

CO-CHAIR THRAEN: Feasibility.

MEMBER ALEXANDER: This is generated

or collected during provision of care, some of it electronically.

The NHSN surveillance protocols, definitions and data collection methods have been used across multiple settings. The collection methods may vary between facilities.

There are no fees but you have to be enrolled in NHSN to participate.

CO-CHAIR THRAEN: Comments or questions. Yanling.

MEMBER YU: Thank you. I know that in Washington State there are sometimes they have to -- our state advisory committee on hospital-acquired infection, they have to do inspection at a hospital, try to determine whether there's under-reporting, over-reporting and reporting error basically.

I'm just wondering whether there's any thought given into it about how those measurements reporting error or under-reporting problem.

And also, our state doesn't like SR. We have our own index. So that means there might

be something, differences interpretation of the infection documentation.

Do you have any thoughts on that?

DR. POLLOCK: Yes, very good questions. I think clearly with respect to your first comment and question there's plenty of room to improve the comprehensiveness, thoroughness of the post-discharge surveillance.

We recognize that this is one of the most important and challenging areas in all of healthcare-associated infection surveillance.

We're keenly interested and we're working very closely both with the clinical community of practice and the ambulatory surgery center environment as well as with innovator initiated strategies for connecting with patients in SSI surveillance, greater use of telecommunications and follow-up.

Speaking of Washington there's a group at the University of Washington in Seattle that has innovated a web-based application called M-Power -- I encourage you to look into it -- that enables

patients in the outpatient phase to report data to the practicing surgeon.

So I think the future of patient-generated health data in post-discharge SSI surveillance is very promising.

We have a lot of work to do on that front, but we're looking forward to working in that area very much.

In terms of our use of the standardized infection ratio we have found it to be an important way to summarize HAI data across the board.

It really is a now very widely used summary metric not only by CDC in its own reports, but also by state health departments in their reports.

Thirty-four states require use of NHSN. Almost all of those states have some element of SSI reporting requirement and are making extensive use of the standardized infection ratio when they report their data publicly.

We also report on behalf of facilities to CMS both surgical site infection data as well

as other HAI data summarized using the SIR.

So, we think it has an important versatility that enables us to provide risk adjustment in a single summary measure that is of value for prevention purposes as well as for quality measurement purposes.

We by no means claim that this is the sole way that data can be summarized usefully. We just have found it to be an effective way of conveying the data and providing guidance both for prevention and quality measurement purposes.

CO-CHAIR THRAEN: Tracy and then Lillee.

MEMBER WANG: I'm just curious. Is the ASCs mandated to use the NHSN? And if not what percent of them are using this database, and are they submitting data consistently and regularly?

DR. POLLOCK: Could you repeat the question again?

MEMBER WANG: Yes. Is it mandatory for the ASCs to send that data to NHSN?

DR. POLLOCK: Okay. So, there are six

or seven states that do require ASCs in their jurisdiction to report SSI procedure and SSI outcome statistics to NHSN. That's why we have a relatively rich database already of ASC SSI data. Colorado is one of the states that has a mandate.

So there are states, but most at this juncture do not require ASCs in their jurisdiction to report to NHSN.

CO-CHAIR THRAEN: Just in terms of the state-based movement in health-associated infections, what Utah has chosen to do is to wait until CMS mandates it and then they mandate it. That's their strategy. I should say our strategy.

Any other questions or comments? Let's vote.

MS. QUINNONEZ: Voting is now open for the feasibility of measure 3025.

Option 1, high. Option 2, moderate. Option 3, low. Option 4, insufficient.

All votes are in and voting is now closed.

For the feasibility of measure 3025 15

percent voted high, 80 percent voted moderate, 5 percent voted low and zero percent insufficient.

CO-CHAIR THRAEN: All right. Usability.

MEMBER ALEXANDER: This is currently being reported to NHSN by those states which he mentioned earlier.

Also, the Colorado Department of Public Health Patient Safety Program is using this for a reporting quality measure.

CO-CHAIR THRAEN: I think that also in the report here the plan is to use it for accountability in the future as well.

So I think right now unless CMS dictates it CDC actually often doesn't have that kind of authority to dictate that this should be used, but CMS does because of the Medicare reimbursement piece and Medicaid.

So now it's still state by state, but the more pressure that's put on to say here's a national measure and we want to be able to evaluate like you were indicating earlier that a lot of the

surgery and a lot of the procedures are going to the outpatient side it actually gives some strength to states to say we want to move in that direction.

Shall we vote? All right, Lisa.

MEMBER MCGIFFERT: I would just reinforce that, that there's -- we work on these issues in a number of states and it is the Wild West. The ASCs will not get all or maybe even most of the surgical centers that are operating out there.

So I don't know how many we'll capture, but I definitely am hearing things from CMS, interest in this.

And from the consumer perspective more and more people are using these facilities and there really isn't any information out there to compare how these do versus a hospitalization. And I think it's really good to see this kind of measure come forward.

CO-CHAIR THRAEN: Ed.

CO-CHAIR SEPTIMUS: Dan, you can correct me on this, but those of us who worked on the HHS Action Plan part 2, in fact ASCs were

included.

And one of the recommendations that came out of our meeting was reporting into NHSN. As you know, we got somewhat of a pushback from ASCs in terms of the timeline, legitimately so because none of them are really equipped to do this.

But I think those of us who worked on that plan, the vision was that eventually they would come online. Is that what your understanding is also?

DR. POLLOCK: Yes. I think that you've summarized it very well, Ed.

It is a priority. There's a lot of work to do to enable ASCs across the board to report.

But we've got I think a reasonable amount of field experience already with the states that have mandated.

And I think part of that experience is we want to make sure in the SSIs that we bring under surveillance and use for quality measurement that we're picking the right procedures.

And the tradeoff between the burden of

reporting and the yield in terms of data for quality measurement and improvement makes it a worthwhile endeavor.

CO-CHAIR SEPTIMUS: And just to speak for it I think we all as has been stated that the oversight in ASCs are not the same as in hospitals.

And by making them more accountable, unfortunately the only way you're going to move the needle is to have a mechanism in place like this to hold them accountable for surgical site procedures that make sense will move us closer to that goal.

DR. POLLOCK: Agree completely.

CO-CHAIR THRAEN: Vote.

MS. QUINNONEZ: Voting is now open for the usability and use of measure 3025.

Option 1, high. Option 2, moderate. Option 3, low. And option 4, insufficient information.

All votes are in and voting is now closed.

For the usability and use of measure

3025 63 percent voted high, 67 percent voted moderate, zero percent for low and zero percent for insufficient information.

CO-CHAIR THRAEN: And then finally suitability for endorsement.

MS. QUINNONEZ: Voting is now open for the overall suitability for endorsement of measure 3025. Option 1, yes. Option 2, no.

All votes are in and voting is now closed. For the overall suitability for endorsement of measure 3025 100 percent voted yes.

CO-CHAIR SEPTIMUS: We're far too easy on you, Dan.

DR. POLLOCK: Well, we appreciate that.

CO-CHAIR THRAEN: All right. We're moving forward into 0450 PSI number 12 from AHRQ. Would the measure developers who are here please join us?

MR. LYZENGA: I should note that this is the second of our two maintenance measures we're considering during this cycle.

So this one is eligible to have us forego the discussion and vote on evidence and reliability if you so choose.

CO-CHAIR SEPTIMUS: Michelle, you're back on, right?

MEMBER SCHREIBER: I am back on, thank you.

CO-CHAIR SEPTIMUS: Okay, great. Thanks.

CO-CHAIR THRAEN: And then, Jason, you're the lead on this one. So we'll turn it over to the developers and then Jason will take over from there.

CO-CHAIR SEPTIMUS: We made Jason the PSI king. Or czar.

DR. PETERSEN: Thank you very much. This is Pam Owens. I am the lead of the AHRQ quality indicators and I apologize that I cannot be there in person.

AHRQ very much appreciates the opportunity to have two of the AHRQ quality indicators reviewed in today's meeting.

PSI 12, Perioperative Pulmonary Embolism or Deep Vein Thrombosis Rate, and PSI 9, Perioperative Hemorrhage or Hematoma Rate.

Since I'm not there in person I do actually believe that Dr. Patrick Romano from UC Davis who is there will do an exceptional job at representing the indicators.

Patrick is the clinical lead of the AHRQ QI contractor team that's led by Stanford University. And both Patrick and I are available to answer any questions although I'm sure you all know Dr. Romano and he is superb.

Before I turn it over to Patrick I do want to take a few seconds to tell you about a few of the core principles of the AHRQ quality indicator program as I believe these are critical aspects to keep in mind during the review today.

The hallmark of the AHRQ quality indicator development process is the continuous enhancement and refinement of all indicators based on user feedback, review of clinical practice changes, validation studies, empirical testing for

validity and reliability, and input from expert panels such as yourselves at the patient safety committee, and experts on our AHRQ QI standing work group.

For instance, we know that the coding of conditions as present on admission has improved over time. And this is a key element in PSI 9 and PSI 12 specifications.

We continuously conduct validity studies and use these results of the studies to continuously improve the indicators.

And I emphasize these last two points in particular because AHRQ is aware of recent publications such as the one by Winters and colleagues in Medical Care that point to validity concerns with PSI 9 and PSI 12.

We also understand that this article was circulated to you as reviewers prior to the meeting. And I wanted to assure you that we address validity in our submission but we don't specifically address the Winters article.

Unfortunately due to the time lag in

publications meta analyses such as this one must rely on older studies, and authors at times do not realize or integrate in their discussions all of the improvements that have already taken place in the indicator specifications and the ones that are being presented to you today.

In fact, the studies highlighted in that article are the foundational rationale for the improvements that we are showing you.

Now, we're happy to address your specific concerns regarding the article as we move through the review, but I just want to leave the discussion there.

Moreover, I wanted to highlight another key component of the AHRQ quality indicator program and that is the transparency and usability of the indicators.

Not only does AHRQ QI program publicly post all of the technical specifications, but we also provide users with SAS and Windows-based software to calculate their own numerators, denominators observed in risk-adjusted rates using

their own administrative data.

Users are a critical component of the QI program at AHRQ. For example, this month we released an updated AHRQ QI toolkit that can be used by hospitals as a general guide to apply improvement methods in a hospital setting as well as guidance on how to improve specifically the PSIs such as PSI 9 and PSI 12.

So, thank you. I will turn it over to Dr. Romano to provide an overview of each of the indicators and as we go through both he and I are available for questions. Thank you.

DR. ROMANO: Okay, thank you. I think I've met most of you before. I'm Patrick Romano. I'm a general internist, general pediatrician and health services researcher based at UC Davis School of Medicine in Sacramento, California.

And I've worked with AHRQ on the enhancement of the AHRQ patient safety indicators. And my team has actually done a number of the validation studies, published a number of the papers that are cited in the submission.

So, I think that I just wanted to make one comment again related to the Winters paper. And this is just a brief statement that's been approved by the editors and publishers of the journal Medical Care which is, and I quote, "Most of the numerical estimates provided I this paper, the Winters paper, are incorrect due to methodologic errors in their meta analyses. The authors are now in the process of re-analyzing their data and submitting corrected results for publication in Medical Care which they have agreed to do. Until these corrected estimates are publicly available readers cannot rely on the published estimates."

And we can get into more details during the discussion review process.

Obviously we start with PSI 12 which is an outcome measure focused on in-hospital venous thromboembolism among surgical patients.

And it's part of a spectrum of quality measures that focus on surgical complications.

CO-CHAIR THRAEN: Jason.

CO-CHAIR SEPTIMUS: Jason, speak into the mike, please.

MEMBER ADELMAN: I just want to start by saying I have incredible appreciation for Dr. Romano and AHRQ and the PSI measures.

I have some issues with them, but I think overall those issues can be worked out, and they're very important for patient safety.

Some of my concerns are big picture concerns. Like for example, we heard that this Winters paper that has methodologic issues, but it also -- I'm sorry, I forgot the name of the woman on the phone from AHRQ. What's her name?

DR. ROMANO: Pam Owens.

MEMBER ADELMAN: Pam Owens. Pam, right, that's right. So Pam mentioned a different point, besides that there's methodological issues Pam mentioned that it looked at data that predates some change in present on admission.

And what I don't understand is that I believe -- both of the PSIs we're looking at today, they talk about ICD-9 and ICD-10.

And people can correct me if I'm wrong, but I think we now live in an ICD-10 world. Like almost all of the validation in both papers are ICD-9 data.

And the measure developers spend a lot of effort to show that any validation stuff pre the POA stuff is -- shouldn't really be counted because we fixed it. And we should look at the new ICD-9 stuff. But now it's all ICD-10.

And then an intellectual argument can be made that ICD-10 is just better, there's many more measures and it's more narrow.

But even AHRQ's language in the application sort of says there are some confusing things about ICD-10 and we still have to work it out.

And I sent around just before an article that said ICD-10 may have some issues.

So, in fact, the measure as it will be applied will be applied to a world of ICD-10. All of this validation stuff that we're going to look at, and please correct me if I'm wrong, will have

been done with ICD-9 data, at least the data that will be presented on the screen, and I'm not even sure that that's like appropriate to judge. I could be wrong and so forgive me.

It's also, it's incredibly difficult. There's hundreds of pages if you look at the applications and the supplements, and there's no like -- an exclusion criteria could have been removed from beforehand till now and it's impossible for us to even know or even scrutinize.

It would be incredibly helpful if from this point on going forward for all PSI measures it uses actually ICD-10 data. And anytime anything is changed have a section that says by the way, we removed these five exclusion criteria and added these six. And we can talk about it and scrutinize it.

But it's just this overwhelming amount of information and data that makes it almost paralyzing. But I am concerned about the ICD-10.

I have another issue which I want to bring up which is that, you know, the last time we

talked about PSIs I had mentioned that UHC had this information out about how to adjudicate PSIs.

And I mentioned it and Dr. Romano suggests we know about it. In fact AHRQ has a toolkit that shows how to do it. And I didn't know that.

So I went back and looked, and in fact they do, of course. It tells you exactly how to craft an email to somebody to say we found this PSI, and it may or may not be accurate, could you please clarify.

People are reporting -- people, our colleagues around the table are reporting that they've done it and they've flipped 20 to 30 to 40 percent PSIs.

So even like this Winters paper where we're questioning the methodology, for this particular PSI, the one we're talking about now, the positive predictive value is a little less than 80 percent.

In the toolkit that was mentioned that was just released and I sent there's something from

Harborview. And they've used the AHRQ approved adjudication method.

And they proudly say that they've flipped 20 percent of the DVTs that were false positive. They went and corrected them. Which is about what the Winters paper said.

And what's most troubling to me is that I believe we now live in a world of hospitals that can afford or have the knowledge to adjudicate and review and reduce the PSIs and those that don't.

If you're wealthier and you care about U.S. News and World Report where they use PSIs, or you care about the HAC penalty then you'll have somebody who will do that process that AHRQ laid out.

But if you don't have the money or the sophistication then you won't. And we sort of need Bernie Sanders to come here and defend the little hospitals, the ones that can't do it. Because we live in two worlds now.

And it makes me wonder if, you know, Dr. Romano said there is this thing, why isn't that part

of the measure?

We just had NHSN. They do have people adjudicate. You had a positive culture and a fever and then you review and you get it.

If that was part of the measure then all this positive predictive value that we talked about, everything would all go away.

And of course the big problem is that it's resources and manpower. And you know, most of our hospitals have NHSN people. But now only some of us have PSI people. And so now we live in a two standard world.

So I had a couple of issues there. We're not really using ICD-10 issues. We live in a world with those that adjudicate and those that don't.

So, we should go through each and every step. But even some of the data we're looking at seems maybe not even the right data.

CO-CHAIR THRAEN: So, because it's maintenance we wouldn't have to go through the reliability and validity process. But because the

source data structure mechanism and definitions have changed, ICD-10, you're recommending what?

MEMBER ADELMAN: Well, I'm -- I don't know. I reviewed the application that describes validity, reliability, and references articles, and provides data for stuff that's not in the world that we live in.

And so as I said an intellectual argument can be made that ICD-10 is just better, but I'll give you an example.

CO-CHAIR THRAEN: Well, before you do that I want to ask the developers. So, in preparation for this presentation have you used the PSI using ICD-10 data, source data, and if so what were your findings, and can you reconcile this concern?

DR. ROMANO: Well, I think as everybody knows we only have nine months now of experience with ICD-10 data in the United States. So I don't think we're alone.

I think all the measure developers that developed and implemented measures using ICD-10

coded data are all in this position of really just getting started on the process of re-validating our measures using ICD-10 data.

There is perhaps one relevant study which was cited from Canada because of course Canada implemented ICD-10 a number of years before us.

And so if you refer to -- and I realize that the submission is quite long and detailed -- but anyway page 22, second paragraph, near the bottom of the second paragraph.

So, Quan et al., sampled patients with PSI events from three Calgary hospitals, reported a PPV for PSI 12 of 90 percent. And that's again just from Canada. We don't know whether that experience will translate to the U.S. or not, but obviously it's an important issue that I think AHRQ will be prioritizing in the coming year.

CO-CHAIR THRAEN: Other questions or concerns about the evidence? Go ahead Yanling and then Leslie.

MR. LYZENGA: I should note that we're

kind of talking about validity right now. So maybe we should start out with the question of evidence and whether there is a rationale supporting the relationship of this health outcome to at least one healthcare structure, process, or service.

And I would remind you that as a maintenance measure we may forego this vote if we so choose.

CO-CHAIR THRAEN: Okay, Leslie, go ahead.

MEMBER SCHULTZ: Just a quick question. Is this version 5.X or is it version 6?

DR. ROMANO: What we're bringing to NQF now is version 6 which is the version that was just released.

And to Dr. Adelman's point, the only real difference between version 6 and version 5 aside from version 6 has a more sophisticated risk adjustment model if you will that accounts for more patient characteristics.

But the other difference is that the specification of the measure now is limited to

patients with so-called proximal deep vein thromboses or pulmonary emboli. So we've removed isolated calf vein thrombi from the definition.

There's a concern about variation in surveillance practices across hospitals that might be influencing that.

CO-CHAIR THRAEN: Yanling and then Albert and then Ed.

MEMBER YU: This may be a dumb question. From my world when I do the study if we do two different types of measurements we look at the error to estimate what the error power is and so you know you have.

Then when you move to ICD-10 you know what the biases or the uncertainty would introduce with new codes.

So, has anyone just by educating myself, has anyone done a study, same set of observations, but analyzed using ICD code 9. And then with the same set of records to look at ICD-10. And then you look at what's the difference. What the uncertainty, what the error power would be.

DR. ROMANO: Yes. The Quan study from Canada that I cited was an example of that. So they did simultaneous ICD-9 and ICD-10 coding.

There's individual hospitals and hospital systems have done a bit of this work. And perhaps some of you may have experience with it in your own systems.

Because of course there was a period during which the implementation of ICD-10 was postponed for a year and then another year.

And so during that period a lot of our hospitals were doing training and sort of making sure that the coding was consistent that our coding teams were doing between I9 and I10.

But virtually none of that work has come into the peer reviewed literature because it's almost all been for quality improvement within our hospital systems. I don't know, others may have experience on that.

MEMBER YU: We have no really the data to really say what its uncertainty is.

DR. ROMANO: Well, all I can say is that

the basic structure of the codes, of the diagnosis codes, is very similar.

So, it's not a fundamentally different structure. There are some slight differences in indexing which we're actually pursuing with the ICD-10-CM Coordination and Maintenance Committee to clarify, for example, for peroneal vein thrombi.

So, there are some slight differences, but in general the basic structure of the codes is very similar.

And I think that's the feedback we've heard from the field as well, that people haven't -- they're not finding anything that's very different.

If you look again at early tracking from some of the systems that we're a part of where there's been early data that's available from ICD-10 we're not seeing dramatic changes in the rates of the indicators.

There's been a general downward trend over time, but we're not seeing a sudden change as of October 1, 2015.

CO-CHAIR THRAEN: Albert, then Susan, then Lisa.

CO-CHAIR SEPTIMUS: I just want to say if it's okay with the committee I don't think we're having problems with evidence and gap. And we're all talking about reliability and validity.

Can we just right now just vote to pass that and then go into a discussion of reliability and validity? Because that's what we're all talking about. So let's move past those two elements.

MEMBER WU: I was going to talk about evidence gap.

CO-CHAIR SEPTIMUS: Oh okay, then I'm sorry.

MEMBER WU: But you short-circuited me. I was actually just going to say let's move along.

I think that we do not need to spend time on this. Since the last time we reviewed this there have been even more studies that show that we can do a lot to reduce thromboembolism. So I

would like to get to a vote.

CO-CHAIR THRAEN: Let's vote. Ed spoke.

MS. QUINNONEZ: We are now voting on measure 0450 Perioperative Pulmonary Embolism or Deep Vein Thrombosis. Since this is a maintenance measure we won't vote on evidence. We don't need to vote on evidence.

CO-CHAIR THRAEN: Okay, so then we'll move forward.

MR. LYZENGA: Is there anybody who would like to take a vote on evidence? Or are we all comfortable?

CO-CHAIR THRAEN: No. Move on. Okay, so gap. Albert, did you want to speak to gap?

MEMBER WU: I'd like to vote on this too.

CO-CHAIR THRAEN: Okay. So what did you say, Albert?

CO-CHAIR SEPTIMUS: He says he wants to vote on it too.

CO-CHAIR THRAEN: What about Lisa?

MR. LYZENGA: We do have to vote on this one.

CO-CHAIR THRAEN: Okay, so we're going to vote on performance gap. We do have to vote on performance gap. So let's call for the vote.

MS. QUINNONEZ: We are now voting on performance gaps for measure 0450.

Option number 1, high. Option number 2, moderate. Option number 3, low. And option number 4, insufficient.

All votes are in and voting is now closed. For performance gap of measure 0450 39 percent voted high, 61 percent voted moderate, zero percent for low and zero percent for insufficient.

CO-CHAIR THRAEN: All right. Let's go to reliability. Who wants to speak on reliability? We've done some discussion about ICD-9 versus ICD-10. Go ahead, Lisa.

MEMBER MCGIFFERT: Well, I did want to address that specifically.

I mean, the reality is that everybody's using ICD-10 finally. And we know that there are

problems with it. But I think it's really important for us not to make any decisions to take these measures off the table because this is the world we live in.

And I believe that ICD-10 is enhancing. I'm sure that there's enhancing information that we have with risk adjustment and everything. And I'm sure there are lots of issues with implementing it in hospitals.

But I think we've seen several other measures that relied on ICD-10 codes in the last day or so, I may be wrong, and this issue didn't come up.

So I think it's just going to be a regular issue as hospitals get more used to using ICD-10.

CO-CHAIR THRAEN: Jason.

MEMBER ADELMAN: I agree with that. Even though I brought it up I didn't mean to say, you know, we just switched to ICD-10, and we don't have the data, and so don't move forward with the measure.

There's a lot of good to all these PSIs.

The part that I don't understand is why not just add -- AHRQ has the measures. They have an official method for reviewing and improving them. They just published abstracts like this very measure from Harborview which I sent to everybody. They showed if you review you can decrease your errors by 21 percent.

And we, like I'm at New York Presbyterian. It's a very big hospital. We have many, many, many people who do the HAIs and only one FTE that does all the PSIs so that we get the benefit of this.

So even though it is a resource it is not a major resource. If AHRQ is formally recommending to review this very measure and give us an example why not just add one little step.

And by the way, that will deal with the ICD-9 ICD-10 issue too because any mistakes or issues a human will review and catch and correct.

And then it'll be perfectly accurate. No more positive predictive values of 60, 70, 80

percent. And then we sit and argue 80 percent is good enough and should be measured but 70 is, well, or 60. You know, we can use all these measures.

I just don't understand why not formalize what you already officially recommend.

CO-CHAIR THRAEN: Susan?

MEMBER MOFFATT-BRUCE: I'm going to help Patrick out here a little bit.

So first of all, Calgary is a university. I may be the only Canadian in the room, but that's an academic medical center. So those patients would be very similar to what we see here.

We also shadowed using the AHRQ software because we have a PSI process that's been in place for two and a half years using ICD-9 and ICD-10. Not much difference.

And in fact, maybe even a little easier to use ICD-10 with the software.

I do have two questions though around the validity and the reliability of the exclusion criteria that maybe I can get some -- I mean I

probably spend two to three hours per week on these validating the cases to Jason's point on top of everything else I have to do, and on top of everything else my team does. So these things are time-consuming because there's a lot of noise and it's hard to interpret them.

In particular, the site of the thrombus. Patrick, you said you're now excluding the superficial ones that we don't treat anyway? Because that was a big problem. And so thank you for that.

Secondly, your exclusion says where a procedure for interruption of vena cava occurs before or on the same day as the first operating room procedure.

So I assume that's an IVC filter? Okay. Can that occur in a different admission as compared to this admission when they might develop a DVT?

So for instance, patients that have big cancer operations. They come in. They get an IVC filter. They get all their preoperative workup

ambulatory. Then they get admitted to have their hemipelvectomy or whatever other. So they're doing the right thing. They're getting them all teed up and then they admit them.

So does this mean, this interruption, this procedure, does that have to occur during the admission?

DR. ROMANO: Okay, so to address a couple of these points. Yes. So the indicator software structure is such that it can only use information during the same hospital stay. So yes. So in order to trigger the exclusion the IVC filter would have to be placed during the same hospital stay.

As you know part of the issue here is that some of the filters can be removed, some can't be removed. So it would be very complicated, not necessarily impossibly, but it would be very complicated to design the logic so that we would know if the filter was actually in place at the time of a surgical procedure.

MEMBER MOFFATT-BRUCE: Well, that's a

huge problem. We're trying to look at resources and getting patients optimized preoperatively, prior to coming for the day of surgery, especially in cancer centers.

So those are my comments. I think you've made it a little bit better, but I will -- I mean to Jason's point these things are incredibly complex and we spend a lot of time managing the data rather than improving the outcomes.

DR. ROMANO: I just wanted to speak to this issue about managing the data and so forth.

And I think to me it's not any different from the healthcare associated infections, the NHSN measures where again hospitals need to have systems in place to monitor and manage the accuracy of the data that are reported.

It happens that this is a different data stream. This is a data stream that goes through Medicare or through state health data agencies. But fundamentally it's just another data stream and hospitals are responsible for ensuring the accuracy of that data stream.

Now, there is some opportunity for many hospitals to drop their rates 10, 15, 20 percent by identifying and avoiding false positive reports.

But I would point out that just recently AHRQ reported from the Medicare Patient Safety Monitoring System data from a national sample of hospitalizations that was subjected to a detailed chart review and a review of all the imaging reports. So completely independent of the codes that were submitted.

And from 2010 to 2014 they reported a decrease from 28,000 to 16,000 post-operative VTE events. So a 43 percent reduction.

So, this is -- independent of codes this is a 43 percent actual reduction that presumably reflects the impact of the process improvements that hospitals are making in response to this focus on the problem of venous thromboembolism.

MEMBER MOFFATT-BRUCE: I don't disagree. It has improved patient outcomes for sure.

It is just with some of the noise. And some of the things you have clarified today will help. It's just a lot of inclusion and exclusion, particularly in this one. And there's another favorite of mine but it's not on the docket today.

CO-CHAIR THRAEN: So, before we go on staff need to clarify something.

MS. MUNTHALI: We just want to make sure that we're holding this measure up to the same standards as other measures as it relates to the ICD coding.

And as you know as of October of last year HHS required implementation of ICD-10. So did we. We required that as part of our submissions.

But we also recognized that it would be difficult for developers to test because the test beds wouldn't be there. So we are allowing a few years of lag time, but we are requiring in submissions that developers provide a statement of their intent for the selection of ICD-10 codes. And I haven't looked at your submission carefully,

but that is something we would have to have for the committee to assess your measure.

Also that you include a full listing of the ICD-9 and ICD-10 codes with the definitions and a conversion table if it's applicable. And also that you describe the process used for identifying the ICD-10 codes.

And so, Patrick, if you can assure that we could have this perhaps by the post comment call would that be?

DR. ROMANO: I think that the ICD-10 as well as ICD-9 specifications are in the packet, in the appendix materials to the packet.

And so there is a table showing those codes side by side.

MS. MUNTHALI: So they meet the requirements for the ICD-10 conversion. Thanks.

CO-CHAIR THRAEN: Thank you. Steve and then I think, who is it, Patricia or Lillee? Lillee.

MEMBER LAWLESS: So, real quick about the IVC filter and the prior hospitalization. Is

it in the preexisting conditions coding on that hospitalization or wouldn't it be?

DR. ROMANO: There currently is -- no, that would require what is called a status code in the ICD-10 lingo and there is no status code.

MEMBER MOFFATT-BRUCE: There is none. And actually we're applying for one because it's a huge problem. We cannot -- so that's the issue. You can't detect that the IVC filter is in place. I've learned a lot about coding here in the last couple of days. This is a big issue I think. But we're going to apply for one.

DR. ROMANO: And I think again AHRQ does work closely with the ICD-10-CM Coordination and Maintenance Committee to support or endorse proposals that are helpful for the quality indicators program.

MEMBER QUIGLEY: Thank you. Dr. Romano and Dr. Owen this is Pat Quigley's voice that you're hearing. I know Dr. Romano that you can see me, but Dr. Owen.

My comments are on behalf of Dr.

Kimberly Applegate who was not able to be with us today. She's a pediatric radiologist and director of practice quality improvement in radiology at Emory University.

And she wanted to express concerns in relationship to the measure since this is a maintenance measure that was brought back to us about limiting the data being collected to only discharges, and that it should have extended 30-day post discharge for readmissions related to DVT and PE.

But the other issue that she had is in relationship to the rationale for excluding those patients who come in with spinal cord injury and head injury, that that really could contribute to under-reporting of these adverse events. And wanted to have some rationale for that exclusion. Thank you.

DR. ROMANO: Yes. So, I'll tackle the second question first and then I'll ask Pam to join me on the first question.

So with respect to the exclusion of

traumatic head injuries, spinal cord injuries, and intracranial hemorrhages this was specifically in response to user feedback as well as feedback from this committee in the past.

And it reflects the fact that clinicians are concerned that their options are really very limited in terms of venous thromboembolism prophylaxis in patients who have acute head trauma, acute spinal cord trauma.

These patients may have clinical contraindications to antithrombotic therapy. And in some cases therefore clinicians have to accept a higher risk of VTE because the risk of hemorrhage would be so catastrophic, or worsening a hemorrhage in this kind of a closed location, in intracranial or in the spinal canal.

So that was the rationale. So it was basically specifically in response to user feedback as well as feedback from NQF stakeholders.

With respect to the readmission question we do acknowledge that some events occur after discharge from the hospital and that many of

these post discharge events may be preventable.

It obviously then becomes a mixture of an ambulatory measure and a hospital measure because the key prevention opportunity then is related to the continuation and the management of thromboembolism prophylaxis after discharge from the hospital.

So as we look forward to the future and really trying to encourage better coordination of care, better handoffs, better integration of care between inpatient and outpatient settings I think we'd agree that these kinds of measures that cut across settings of care would be valuable.

Pam, do you want to address AHRQ's perspectives on this question with respect to the QI program?

DR. PETERSEN: Well, in terms of from a QI perspective as you know from our submissions at the moment it relies on all payer data from the Healthcare Cost and Utilization Project. That is a discharge database.

It does not have -- we include treatment

EC visits in that database as well as ambulatory surgery visits and hospital affiliated settings.

That being said we have discussed internally the need to think about expanding in terms of looking at the entire episode and looking in the outpatient arena.

It would limit. We would not be able to do it all payer, but it is definitely worth some explorations and actually some growth areas that we've already talked about internally.

So thank you very much for putting that as a suggestion.

MEMBER QUIGLEY: Thank you.

MEMBER ALEXANDER: My ask is similar. It would be so interesting to look at readmissions and be able to tie that back into a prior surgical procedure. And that would be a metric that I think would be greatly of value if you could develop that.

DR. PETERSEN: To that end we do do quite a bit of readmissions work. There are a subset of the states, there's 46 states that I believe participate in the Healthcare Cost and

Utilization Project.

A subset of them have unique identifiers that allow us to link hospitalization. Certainly it is an area ripe for continued study to look at it from a readmission standpoint.

We don't get the outpatient visits for instance and the physician visit arena, but we could link some hospitalizations for a subset. So we'll look at it.

MEMBER ALEXANDER: And even if you could pull, and this might be too hard to gather, but if you can get emergency room admissions that would be helpful.

DR. PETERSEN: Yes, and that's definitely -- we could include that as well.

CO-CHAIR THRAEN: All right. Are there any other questions about reliability? Shall we vote?

MS. QUINNONEZ: Voting is now open for the reliability of measure 0450.

CO-CHAIR SEPTIMUS: Just keep in mind what Lillee said about the issues between ICD-9 and

ICD-10.

DR. ROMANO: I'd also like to call people's attention to in the technical submission materials actually we do specifically address the historical trending of the evidence over time.

And so there is a paragraph in there specifically about the studies that followed the advent of POA coding. And specifically one study showing a PPV of 99 percent, others showing a PPV of 81 percent following POA coding.

And additional single center studies with PPVs in the range of 88 percent and 93 percent.

So I think that it's just important to note that as you focus on the more recent studies that reflect the advent of POA coding as well as changes with more specific ICD-9-CM codes that the positive predictive values have increased as you would expect.

MEMBER WU: Though, Patrick, we're talking about reliability now.

DR. ROMANO: Understood. Nobody's really talked about reliability here.

MEMBER WU: We've been talking about validity.

DR. ROMANO: There is reliability data in the submission.

CO-CHAIR SEPTIMUS: We always stay on point, Patrick. Let's vote.

MS. QUINNONEZ: We are now voting on the reliability of measure 0450.

Option 1, high. Option 2, moderate. Option 3, low. And option 4, insufficient.

All votes are in and voting is now closed.

For the reliability of measure 0450 11 percent voted high, 78 percent voted moderate, 11 percent voted low and zero for insufficient.

CO-CHAIR SEPTIMUS: Okay. So just a small break here. Iona unfortunately has to catch a flight so she's going to leave in the next few minutes, but I wanted to take personal pride in thanking her for working with me and of course co-leading this committee. But I think she deserves a round of applause.

(Applause)

CO-CHAIR SEPTIMUS: Okay, now we're going to validity which I think we have sort of talked about extensively. But there's always one in the crowd, Missy. No, no, I'm kidding.

MEMBER DANFORTH: I was waiting for the appropriate time.

I have a question. No, it's not a question, it's a clarification.

I noticed that pregnancy is an exclusion. Does that mean women undergoing C-sections are excluded?

DR. ROMANO: Yes, that's correct.

CO-CHAIR SEPTIMUS: Let's vote since we've talked about validity.

MS. QUINNONEZ: Voting is now open for the validity of measure 0450.

Option 1, high. Option 2, moderate. Option 3, low. And option 4, insufficient.

All votes are in and voting is now closed. For the validity of measure 0450 18 percent voted high, 76 percent voted moderate, 6

percent voted low and zero percent for insufficient.

CO-CHAIR SEPTIMUS: Okay, so now we're up to usability. Jason, any comments on usability? I'm sorry, feasibility. Any comments on feasibility?

MEMBER ADELMAN: I think the measure is both feasible and usable.

CO-CHAIR SEPTIMUS: So if there's no comment let's take two votes in a row. Is that okay? All right, so let's do it. This may be a first.

MS. QUINNONEZ: Voting is now open for the feasibility of measure 0450.

Option 1, high. Option 2, moderate. Option 3, low. And option 4, insufficient.

All votes are in and voting is now closed. For the feasibility of measure 0450 76 percent voted high, 24 percent voted moderate, zero percent for low and zero percent for insufficient.

CO-CHAIR SEPTIMUS: Okay, now we'll go to usability.

MS. QUINNONEZ: Voting is now open for the usability and use of measure 0450.

Option number 1, high. Option number 2, moderate. Option number 3, low. And option number 4, insufficient information.

Can we have you resubmit your votes one more time please to make sure capture all the votes?

All votes are in and voting is now closed. For the usability and use of measure 0450 71 percent voted high, 29 percent voted moderate, zero percent for low and zero percent for insufficient information.

CO-CHAIR SEPTIMUS: So we're up to the last question.

MS. QUINNONEZ: Voting is now open for the overall suitability for endorsement of measure 0450.

Option 1, yes. Option 2, no.

CO-CHAIR SEPTIMUS: And this is re-endorsement actually, right?

MS. QUINNONEZ: Re-endorsement, thank you.

All votes are in and voting is now closed. For the overall suitability for endorsement of measure 0450 100 percent voted yes.

CO-CHAIR SEPTIMUS: Okay. Now we're up to the last measure 2909 which is Perioperative Hemorrhage or Hematoma Rates. It's PSI 9 also from AHRQ. So, Patrick, I guess I'll turn it over to you.

DR. ROMANO: Yes, so very briefly this is just another one of our post-operative complication measures.

This measure focuses on post-operative hemorrhage or hematoma and it requires a diagnosis of the same along with a return visit to the operating room, or a follow-up procedure.

And the effort with this is to identify a subset of hemorrhages or hematomas that are associated with some kind of intervention or some need for follow-up care.

This measure again has been used and it's really quite appropriate that it's considered

together with PSI 12 because we've heard from users some concerns that as they try to aggressively prevent thromboses they may be causing more hemorrhages, or conversely if they try to prevent hemorrhages they may be causing more thromboses.

And so it's appropriate that the committee consider these two measures together because they are to some extent designed to assess two sides of the coin if you will related to two different outcomes.

And obviously our aim in the hospital business and healthcare business is to try to minimize both. But in some cases there may be some tradeoff.

CO-CHAIR SEPTIMUS: And Jason has graciously consented since Linda's not here his tremendous knowledge in this space. So Jason, let's go through. Start with the evidence and gap if you feel that we need to re-discuss this.

MEMBER ADELMAN: Right. As far as evidence and gap I think that there is evidence that this is an important issue and there is certainly

room for improvement.

I have very brief comments about the other aspects, but we could just vote on that if you like and then move on.

CO-CHAIR SEPTIMUS: Well, does it get to reliability and validity?

MEMBER ADELMAN: Just that all of the same things that we already discussed we don't have to discuss again.

The only thing I would say is this measure more so than the one we just discussed I've -- despite that there are studies that show varying PPVs and that in setting up the process for adjudicating PSIs at New York Presbyterian I spoke to some of my colleagues around the table and many people around the country, at the Brigham, and Mount Sinai.

And this one much more so than the others people are flipping 30-40 percent. So the reported positive predictive value could be 80-90 percent, but people are flipping a lot more.

And it has to do with exclusion criteria

of platelet deficiencies that are caused by Plavix.

So somebody goes for a cardiac cath. They're having a cardiac incident. They're on something that causes platelet deficiencies by choice which is an exclusion criteria. And then there's clarification later on that the coders couldn't really pick up because they can't make diagnoses.

So I suspect that the true positive predictive value is worse here, and the need for reviewing is more important here. And those of us at hospitals that are doing it are benefitting. Those that aren't, aren't benefitting.

That's the only thing I wanted to add above what we already discussed.

CO-CHAIR SEPTIMUS: Thank you very much, Jason.

So, do we want to vote? Steve.

MEMBER LAWLESS: Actually, just real quickly. Are you seeing hospitals that actually have high bleed rates and are doing the prophylaxis well? I mean, I always worry the

counter-balancing. If you put it into quadrants what is the data showing?

DR. ROMANO: I would need a couple of minutes to pull up details, but I will say that in general when we -- so, you may recall we discussed the PSI 90 which is the composite PSI I think last year.

And so in the construction of that PSI we included both PSI 9 and PSI 12. And we found in that process actually that there was a positive overall correlation. So the hospitals that had higher PSI 12 rates in general also tended to have a little bit higher PSI 9 rates.

So in fact the hypothesis was not supported in that analytic work. But of course it may be that if you focused on particular subsets of surgical patients that you might find that negative correlation. But overall we found a positive correlation consistent with their both being quality metrics.

CO-CHAIR SEPTIMUS: But positive correlation meaning they're opposite. So you want

to have a negative correlation. Sorry, put it this way. In the process of those hospitals that are successful at lowering thromboembolism, are they having a higher bleed rate?

DR. ROMANO: Well, one would be concerned that they might be over-treating for thromboembolism and we did not find evidence of that.

MEMBER ALEXANDER: I just have to say I think this is a much improved measure now that you've added in the re-operation or intervention.

Anyone can bleed and I think today, Jason's point, we've got people on so much blood thinner right now and there are all sorts of interventions we're doing before surgery to try and decrease bleeding. Every patient in my hospital is on something when they come in. And so people can bleed.

What this to me does as a surgeon which I think is so far superior is anyone can bleed. It's my job in the operating room to control it as best I can. Whether that is pharmacologically, or

whether that's electro with a Bovie or something like that.

This picks up the people I don't control, that I don't identify that it's an issue and can't control. And I think it's a far superior measure.

CO-CHAIR SEPTIMUS: Okay, so do we want to vote on the evidence?

DR. ROMANO: Could I also address Dr. Adelman's other point?

CO-CHAIR SEPTIMUS: Please.

DR. ROMANO: So, I think Dr. Adelman raised a very important point which is about the exclusions and the possibility that different hospitals may sort of game the exclusions or manipulate them.

And this is a real concern. And we're -- certainly we're always monitoring feedback from the user community and attempting to respond to it.

I will say that the exclusion here is designed specifically to capture and therefore exclude patients with congenital factor

deficiencies, patients with specific syndromes, thrombocytopenic syndromes, platelet defect syndromes.

But it does not exclude patients who have medication-induced clotting disorders. So that's really important.

Now, it may be that some people are lying about the codes and they're using these codes. But the guidance from the Coordination & Maintenance Committee, from coding clinics has been very clear that these codes that are used for exclusions are not the right codes to use for medication-induced clotting problems.

CO-CHAIR SEPTIMUS: Okay, so the question is do we want to vote on this? Yes, Jason.

MEMBER ADELMAN: My point was not that people were manipulating because honestly I actually didn't know. I guess there's a level of sophistication to understand how to do a proper adjudication.

And Dr. Romano, you made a point earlier about the similarities of this with the HAIs.

But with the HAIs NHSN described every step of the process, like exactly how to define catheter days, and temperatures within catheter days.

And here we have recommendations to review. I mean, if you look at the requirements it's written under exclusion. It's just written platelet defects I think, or qualitative platelet defects, that's what it says. Qualitative platelet defects.

And so it doesn't say hereditary qualitative and it doesn't say drug-induced. It just says qualitative. And how is anyone supposed to -- like I was honestly looking and trying to understand.

And I don't go to code clinics. I don't know that those guidelines are only for -- and so I still see the value and would ask that you just consider to formalizing the review process and then give better and better guidelines to how to do that.

Because I think many people are well-intended and think that giving Plavix causes

a true qualitative platelet defect. If it said non-drug induced then I would think that they're being egregious. But anyway, I think we should formalize and put part of the measure all of the PSI measures in the review. And it would be much better.

CO-CHAIR SEPTIMUS: Thank you. Okay, so are we ready -- do we want to vote on evidence? I hear yes so let's vote on the evidence.

MS. QUINNONEZ: We are now voting on measure 2909 Perioperative Hemorrhage or Hematoma Rate PSI 09. Voting is now open for evidence.

Option number 1, yes. Option number 2, no.

All votes are in and voting is now closed. For the evidence of measure 2909 100 percent voted yes.

CO-CHAIR SEPTIMUS: Okay. Lillee.

MEMBER GELINAS: Just a point of order and maybe for the NQF staff as we lose people what is the finite quorum that we fall below? I just need clarification from a true numbers standpoint

for the voting numbers. What is a full quorum?

MS. QUINNONEZ: So, we need 13 committee members to proceed. So we're tracking that. Right now I think we're at 16.

MEMBER GELINAS: Thank you.

CO-CHAIR SEPTIMUS: Michelle, you're still on the line, right? Listen, this is the last measure. If we stay focused I think we can certainly finish before any other people have to leave. So we're well above a quorum.

So let's stay focused and talk about gap. Do we need any discussion on the gap? Then let's vote.

MS. QUINNONEZ: We are now voting on performance gap of measure 2909.

Option 1, high. Option 2, moderate. Option 3, low. And option 4, insufficient.

We should still have 16. Would you please resubmit your votes? They're coming in now.

All votes are in, thank you, and voting is now closed.

For performance gaps of measure 2909 40 percent voted high, 60 percent voted moderate, zero percent for low and zero percent for insufficient.

CO-CHAIR SEPTIMUS: Okay. The next one is reliability. Any other discussions? I see no. Then let's vote on reliability.

MS. QUINNONEZ: Voting is now open for the reliability of measure 2909.

Option 1, high. Option 2, moderate. Option 3, low. And option 4, insufficient.

All votes are in and voting is now closed. For the reliability of measure 2909 40 percent voted high, 60 percent voted moderate, zero percent for low and zero percent for insufficient.

CO-CHAIR SEPTIMUS: Okay, the next one then is validity. Any more discussion on validity? Seeing none we'll vote.

MS. QUINNONEZ: Voting is now open for the validity of measure 2909.

Option 1, high. Option 2, moderate. Option 3, low. And option 4, insufficient.

If you could resubmit your votes one

more time, please. Thank you.

All votes are in and voting is now closed. For the validity of measure 2909 33 percent voted high, 67 percent voted moderate, zero percent for low and zero percent for insufficient.

CO-CHAIR SEPTIMUS: I think we're now moving to feasibility. Any discussion on feasibility? Seeing none we'll vote.

MS. QUINNONEZ: Voting is now open for the feasibility of measure 2909.

Option 1, high. Option 2, moderate. Option 3, low. And option 4, insufficient.

Thank you. All votes are in and voting is now closed.

For the feasibility of measure 2909 80 percent voted high, 20 percent voted moderate, zero percent for low and zero percent for insufficient.

CO-CHAIR SEPTIMUS: I think the next one is usability. Any discussion on usability? Jason's ready so we're all ready. Okay, go for it.

MS. QUINNONEZ: Voting is now open for usability and use of measure 2909.

Option 1, high. Option 2, moderate. Option 3, low. And option 4, insufficient information.

All votes are in and voting is now closed. For the usability and use of measure 2909 87 percent voted high, 13 percent voted moderate, zero percent for low and zero percent for insufficient information.

CO-CHAIR SEPTIMUS: And now the drumroll. Is this measure suitable for re-endorsement?

MS. QUINNONEZ: New measure.

CO-CHAIR SEPTIMUS: I almost got through without making a mistake. I apologize.

MS. QUINNONEZ: Voting is now open for the overall suitability for endorsement of measure 2909. Option 1 is yes and option 2 is no.

All votes are in and voting is now closed. For the overall suitability for endorsement of measure 2909 100 percent voted yes.

CO-CHAIR SEPTIMUS: I must say we've been very nice to you, Patrick, this time. But I

think you've heard some really good suggestions from the committee. I think we're all struggling with the issue of ICD-9 and ICD-10. And I think that that's something you should continue to validate over time and update the measure when you think it's appropriate.

I think as Jason well said before the work your team and AHRQ has done in providing these measures and being responsive to this committee. We all know the history of PSI 90. We certainly appreciate your time. And actually coming here in person again.

And also thank Pam Owen. Are you still on the phone, Pam?

DR. PETERSEN: I am. Thank you very, very much for all of your comments.

CO-CHAIR SEPTIMUS: So thank you and I hope that you'll take those comments back and keep them in the queue or the parking lot for future development. So thank you very much.

DR. ROMANO: Thank you.

CO-CHAIR SEPTIMUS: Okay. So, I guess

we can go through next steps pretty quickly that you see on the screen.

We have some business we're going to go back to. We are actually I think doing really well on time and I thank everyone for their focus and their comments.

So, here's the next steps. I think post in-person meeting call I do not believe will be needed. So I'm sure you're all very hurt that you can write that off your calendar and get a couple of hours back.

We'll be drafting the report to NQF members and the public early this fall. And as you know we will have a review and a standing committee call to review those public comments. Then we'll draft the report to the NQF member vote. It then goes to CSAC for review and approval. And then eventually goes to the board for endorsement just before Christmas.

And if you remember there is also another period of appeal. And for those of you who went through the original sepsis measures back four

years ago in fact there was some additional comments even after the endorsement from the board. I don't think any of our measures are going to go through that kind of problem, but that's the timeline.

MR. ANDERSON: So we're actually going to -- remember we had tabled that measure 3000.

MR. LYZENGA: We tabled a decision on that one. So we will hold the post meeting call. There's also one earlier today that we offered the opportunity.

CO-CHAIR SEPTIMUS: Do you think they can get the information back? Okay, I see what you're saying.

MR. LYZENGA: So we will need that call.

CO-CHAIR SEPTIMUS: Well, actually the call is in --

MR. ANDERSON: It's in a week and a half.

CO-CHAIR SEPTIMUS: Do you think -- I mean I'm asking, do you think they're going to come

back that quickly?

MR. LYZENGA: So, post comment then? So we'll have to wait until the post comment call.

CO-CHAIR SEPTIMUS: So I don't think -- they're not going to be ready. At least that was my understanding.

MR. ANDERSON: Fair enough.

MR. LYZENGA: Right.

CO-CHAIR SEPTIMUS: Okay, maybe you ought to hold it, but I doubt we're going to have a call on the eighth. That's probably the best way, to hold it, okay?

So, let's -- before we get into other stuff I think just to make sure we cover this is there any public comment, Operator? Anybody in the room? Public comments on the phone, Operator? Operator? Are you on mute?

OPERATOR: My apologies. Once again to make a public comment please press * then the number 1. And there are no public comments at this time.

CO-CHAIR SEPTIMUS: Thank you very

much, Operator.

So, I think we'll go back to what we skipped over this morning to make sure we got through the measures and talk about -- discuss gaps in measurement. Andrew, do you want to lead that discussion?

MR. LYZENGA: Sure. So thanks, everybody, for staying in. We can maybe revisit this as well on one of the post meeting or post comment calls just to get some input from the rest of the folks who have scattered at this point.

But we did want to talk a little bit about gaps in the portfolio and gaps in measurement around patient safety in general.

NQF, we like to do this in each cycle just to give feedback to developers.

NQF is increasingly looking to get involved as part of our strategic direction in the identification and prioritization of gaps in measurement in general. So this is a good opportunity to get some input and for us to get some feedback from you on that as well.

We also have as we mentioned yesterday I think Helen was going to talk a little bit more about it but she's gone at this point.

But the measure incubator is an opportunity to sort of advance new measures in gap areas and particularly innovative measures that are taking approaches that we haven't necessarily seen before.

So, in any case we would like to get some thoughts from you if you don't mind on gaps in our portfolio of measures in terms of topic areas, types of measures, and also just thoughts on new or promising approaches to measurement. Things that maybe we can think out of the box on. Just your thoughts about where the future of measurement should go in patient safety.

So I'll just open up the floor with that. It looks like we've got a couple of folks already.

MEMBER LAWLESS: Actually, you heard it in a couple of comment themes, interoperability and safety around the lack of interoperability, and

then transitions of care.

We've nailed every process you can find in an area, but if there are processes involving multiple areas that would be pretty -- it sounds novel, but it would be very, very beneficial.

CO-CHAIR SEPTIMUS: Can I ask the NQF staff, I mean interoperability has been a thorn in many of our sides for quite some time. And I think you've had some discussions about this. Is there anything that we can do to sort of get folks together to try to make this actually finally happen? It's been a decade or more since we've been talking about.

I'm not talking about NQF responsibility. I'm talking about how NQF as a non-biased group can pull these folks together.

MR. LYZENGA: I can tell you it's definitely an area of strong interest for us. And we're looking to do more work around interoperability.

I was involved as was Jason Adelman in our HIT safety work. There's a project around

identifying and prioritizing measure concepts related to HIT safety and interoperability came out as one of the highest priorities for trying to ensure the safety and safe use of HIT systems.

And so there is definitely some push to develop some measures and approaches to increasing interoperability that came out of that work.

I would anticipate that we'll be doing more work in that area in the future without question.

CO-CHAIR SEPTIMUS: Steve, did you have any thoughts about -- I mean, I agree transitions of care are probably one of the most dangerous times for patients.

Did you have any -- I mean, you don't have to state it but if you have any thoughts.

MEMBER LAWLESS: Two I'll give you right off is if you look at the Joint Commission data on sentinel events communication handoff is 80 percent of them. And so that's one.

And the other is if you talk about readmissions you can discharge somebody and then

you can accept somebody. You can discharge and then you can accept.

And so the discharging of a patient from the hospital to what happens immediately after discharge, that's a transition.

And we've had that sometimes here around thromboembolic agents and stuff. But those are the two transitions I would say people designing measures around would be great.

CO-CHAIR SEPTIMUS: Thank you very much. Charlotte and then Lisa.

MEMBER ALEXANDER: I have a lot. Sorry. Wrong site surgery in ASCs. We see a high instance of that.

I'd like to pair 2951 with referral for treatment.

I'd like to get some stuff focused around episodes of care. Where we're moving is trying to be able to manage people outside of the hospital. That's where most of healthcare is determined.

An indirect way might be to pull things

like readmissions for diabetes or end stage renal disease because our diabetes readmissions are usually because we're not doing an adequate job with education, or follow-up, or getting them into their doctor, or seeing if they can pay for their meds, or seeing if they've got a home to live in, or food security.

I mean, there's all this stuff we're not fixing and the only way I know we can measure it is when they bounce back into our facilities.

So if we can start getting beyond just what we're doing in the facilities and into what's a reflection of where we're falling down in the rest of healthcare. To me that will start driving people to do a better job in that arena.

CO-CHAIR SEPTIMUS: Great points.

MR. LYZENGA: I should note that we have reviewed a number of readmissions measures for different settings including dialysis facilities, I think some other post-acute care type settings recently and most of those I think have passed through. So we've got some stuff going through the

pipeline.

CO-CHAIR SEPTIMUS: Yes, and obviously one of the things that we're not always aware of is what other committees are reviewing.

But nonetheless it may or may not be covered by another committee. So, Lisa.

MEMBER MCGIFFERT: I just feel like we have a pretty large gap of medical error measures.

And we usually parse them up into little pieces, or we have like PSI 90 that brings some of them together. But we really don't have a sense of what's happening out there and we know there's a lot going on. So I'd like to see some work on that.

And I would like to see a measure that would measure the accuracy of administration and billing data.

CO-CHAIR SEPTIMUS: You know, for some diagnoses there is. And sometimes it correlates and sometimes it doesn't.

Obviously administrative data is very heavily dependent upon physician --

MEMBER MCGIFFERT: Notes?

CO-CHAIR SEPTIMUS: Yes, progress notes. But that's right.

MEMBER MCGIFFERT: So, I have a philosophy that this is the hospital's responsibility as well as the physician's responsibility to be accurate in these records.

And so there may be ways to look at this electronically, but there also may be ways that you would validate it with checking the charts.

I've been doing this work or about 30 years now and for 30 years I've been hearing people say oh, you can't use that data because it's not accurate, and it's only for billing.

And we know that's not true. And I think that the public might benefit from knowing which hospitals are unable to accurately document for billing purposes as well as quality purposes.

CO-CHAIR SEPTIMUS: Lisa started when she was 15. Yanling.

MEMBER YU: I would like to say more measures be encouraged towards ambulatory surgical

center safety and infection.

This one today we reviewed is very refreshing. I hope to see more of those types of measure be encouraged.

Another thing is I would like to see more measurements to incorporate more about public reporting.

Sometimes in those measures when the developers submit it they say usability, they just say we have a plan for public reporting in the future. But you don't know really what their plans are.

I would like to see them actually have more text in there, more meat in there, exactly what they have in mind so that when we look at it. Because we all know public reporting is an important factor when it comes to the measure. So it would be nice to include that.

CO-CHAIR SEPTIMUS: Is that -- let me see if I can. Would you like to see does public reporting impact safety? Is that what you're getting at?

MEMBER YU: Yes. How public reporting would improve the outcome, the quality of the care in terms of that.

So, the bottom line is we want to improve the quality of care and the safety. That's our goal.

MR. LYZENGA: I should also note that the use and usability is -- we'd like to put increasing emphasis on that for maintenance measures particularly.

So when you see a measure come back as a standing committee and they've said something like that, we plan to put this into public reporting, you should certainly hold their feet to the fire and hold them accountable for doing what they have said they were going to do.

MEMBER ARDIZZONE: This might be my own naivete. I don't know what your full portfolio looks like.

But especially at my institution we provide just as much outpatient care as we do inpatient now.

Sixty percent of all our oncology care is outside the hospital. People get chemotherapy outside the hospital. We do bone marrow transplants in patient's homes.

I would love to see reliable, valid outpatient indicators. We do really well in the hospital now, we've got that down to a science. We just struggle to find good, reliable outpatient measures that indicate quality.

CO-CHAIR SEPTIMUS: So I think I hear a little bit of a theme here about outpatient. So we might want to take that one back.

MEMBER GELINAS: When it comes to the world of safety the bucket of workforce measures. I haven't heard that a lot in our work so far.

Individually, at the state level, at the regional level we are showing correlations between nursing turnover and harm, nursing competency and harm. Fatigue and harm.

And we tend to go to the easy stuff. Believe it or not it's easier to measure the financial stuff than the workforce stuff.

But I commend NQF because that whole nursing bucket of work was one of the very first ones that NQF ever did.

I think as we move into value-based purchasing and thinking of care overall, not just care in silos, one of those crosscutting measures has to do with workforce.

Not just nursing, but since nursing is half of the workforce it's a decent place to start for impact.

The whole concept of a balanced scorecard, trying to get to safety is something that I think NQF could do a white paper on, convene an expert panel.

We keep talking about gaps in measurement, but I'd like to perhaps put the concept gaps in wisdom. Gaps in wisdom.

Which is why groups like this work because this is wisdom sourcing. You know, we've heard of crowdsourcing for projects and other types of things, but truly what these panels are are wisdom sourcing.

So, it doesn't necessarily have to be a measure, but it would inform the field around important topics related to safety, or related to quality that I think NQF could have a tremendous contribution towards.

So workforce measures, financial measures tied to workforce measures so we begin to see the balance there.

Causation. It will never be cause and effect, but we can show correlations, I do believe that.

And then that whole world of ambulatory that I think we've talked about.

The Joint Commission just had a panel on ambulatory nursing. And we learned a lot about why certain outpatient centers, ambulatory care centers, physician practices, primary care clinics don't hire registered nurses because then they don't have to meet the Nurse Practice Act requirements.

And so if you don't hire the nurse, or the pharmacist, or the licensed clinician you don't

have to meet the state mandates. That's a workaround that's gaming the system and that's not getting the public what they need in terms of safety. So, thank you.

CO-CHAIR SEPTIMUS: Those are excellent. So I think the workforce issues and nurse turnover rates for those of us who primarily work in the inpatient setting is so very, very important.

And we not only need to sort of figure out how to measure that because as Lillee just said turnover rate is directly related to some outcomes. And it's really difficult to sustain quality programs with a 15-20 plus percent turnover rate.

I think it's personally a very, very high priority on the inpatient side. And there's some opportunities certainly on the outpatient side as well.

So we've had a couple of measures come through looking at level of nursing, RNs versus associate degrees, et cetera, looking at outcomes and I think that's a good first step.

The fatigue issue, I wonder if you'd comment a little about this. Because there's a lifestyle issue that I'm not sure will ever go back. But the shifts are commonly 12 hours for nursing.

And we know that as they get towards the tail end, and tell me if I'm wrong about this, at the tail end of that that, that's when a lot of mistakes can be made.

And I can't remember what hospital in what country this was, I think it was in Europe. They actually went to shorter shifts and they actually got much better employee satisfaction actually.

But then there's that balancing act between lifestyles and getting an extra day off per week versus working a really long shift.

I mean, to me there's no such thing -- I've never come in at change of shift and seen the nurse go home at 7. So I don't know whether you have a comment about the fatigue factor.

And the burnout factor. You know, talk about nurses being in front of the computer and not

being able to interact with patients. All this is a dissatisfier for nurses.

MEMBER GELINAS: There's a growing body of evidence around the direct link between fatigue and error.

And it's not well known I believe, but because of the high nursing turnover today, particularly with new graduate nurses, it can be as high as 40 percent in the United States. New graduate nurses don't stay in their job any longer than two years.

And if it costs between sixty and eighty thousand dollars per nurse to replace them and reorient them we're talking real bucks in the United States with the amount of turnover.

Some of that is related to fatigue, Ed, but I will tell you there's been debate about 12-hour shifts for a long time.

And if you were just to work what you're supposed to, the three 12-hour shifts and then have two off. But to your point it's never just three 12-hour shifts.

It's also a financial factor because nurses are picking up those extra shifts to make money because of the economics of what the average nurse makes.

We talked last night at dinner about traveling nurses and they're going to other countries, experienced nurses, because they make so much more money.

So there is the body of evidence around fatigue and error, and then there is the knowledge about it. So there's the research and the body of evidence, but who knows it and who's applying it? Why should we?

So, the more this whole patient safety arena becomes a public cry, when the public begins to realize how important it is. Could you imagine a patient saying to their nurse how long have you been awake before you try to start that IV?

Or some of those types of consumer issues. It'll be a very interesting day.

But I think at the very least awareness building is where we need to be right now because

the body of evidence around fatigue and error, not just in healthcare but in other industries is pretty robust.

CO-CHAIR SEPTIMUS: And the other thing is that alignment of healthcare professionals at the local level, how we communicate with each other, and is it really, truly a teamwork environment, or is it hierarchical are big dissatisfiers for folks.

So, I think those are great comments. Missy.

MEMBER DANFORTH: Just two comments on gaps and then one request from NQF staff.

So first, we've been doing some work for the past 18 months at Leapfrog around diagnostic error. We assembled a national expert panel about two years ago in person actually at Armstrong Institute and asked them to help us identify the biggest gaps in measurement related to patient safety and they unanimously said diagnostic error.

So we've been doing work with folks like David Newman-Toker, and Mark Graber, and Paul

Epner, but I think that's the big opportunity related to patient safety.

The second one is patient-reported outcomes. We're part of lots of conversations, and meetings, and collaboratives, and groups. There's a lot of interest in patient-reported outcomes, particularly related to patient harm and preventable harm. I mean, this group has talked about it.

Then the request is that I think it would be helpful to get to Laura's comment about the portfolio to have like a summary based on care setting. So hospital, outpatient, ambulatory, dialysis center, physician office, the number of process/outcome/structural measures.

And then just like I don't know where to find that. I mean, I'm on that QPS website probably as much as anyone, but it's really hard to just get summary level.

MR. LYZENGA: Yes, it's tough. We can do that. I was actually doing a little work to do that in the runup to this meeting and didn't have

all of it together. But I can pull it together by a bunch of dimensions, measure type, care setting, level of analysis. And we'll circulate something like that.

MS. MUNTHALI: We're in the process of revamping QPS where we're going to do that for all of the topic areas and subtopic areas in QPS.

It's part of our new strategic plan, to make sure that we can slice and dice the data for all of our measures by all of those levels that folks would want to see.

And so hopefully by the end of fourth quarter, first quarter of next year.

MR. LYZENGA: I should also say we have every intention of pursuing work around diagnostic error too.

And patient-reported outcomes is a major focus of the work in the incubator right now.

CO-CHAIR SEPTIMUS: I want to second diagnostic errors was one of the ones on my list. Oh, yes. Well, you're the great mind, I just follow your lead. I learn from you.

MEMBER MCGIFFERT: One thing about patient-reported outcomes that we've talked about is adding questions on the HCAHPS so that it would be standardized and incorporated in a practice.

CO-CHAIR SEPTIMUS: Good point. Jason? And then Albert. Jason, then Theresa, I'm sorry. Then Albert.

MEMBER ADELMAN: I would love, I don't know if this is possible, but I would love it if NQF could themselves have a composite measure that was simply a harm measure.

And what I mean by that is take the multiple NQF-endorsed measures that come from many different developers, NHSN, AHRQ where the point is really just to communicate the measurable amount of harm that we have.

So for example, there can be at any given hospital, average sized hospital 40 people a year that when they fall they get hurt, and 20 CLABSIs, and 30 CAUTIs, and this number of hemorrhages.

And so out of every 1,000 patients

there's a rate of 50 people being harmed. So it's just a harm scale. So that we can then drive to reduce it.

Now, all the measures I'm talking about are the exact same measures that are in value-based purchasing and HAC penalties, they're all there and they're used as individuals.

But I think it would be easy for the patients and the world to translate. It's a lot of harm per 1,000 patients and we want to see that number go down.

And the reason why I say NQF to build it is because it's really just adding up the harm from all these other developers.

But like let's say I wanted to do it as a researcher and patient safety officer from Columbia. I just can't keep track of where AHRQ is with their measures, and NHSN is.

You're sort of well situated. It's really just adding them all up and they're already endorsed. So it maybe doesn't even need to be reviewed. It's just a different way of doing what

value-based purchasing is doing, but translating it to something that the world can understand.

CO-CHAIR SEPTIMUS: I was really thinking along the way that I think a really great measure developer in this area is someone named Jason. Maybe you could even work with Patrick and do that.

I mean, it's a very good thought. We're all searching for either that measure or that composite that really is the most predictive of harm. And I think that those are good points. Theresa.

MEMBER EDELSTEIN: So, I would like to see -- this won't come as a surprise -- more measures in the post-acute care space, particularly in support of Steve Lawless on transitions of care.

As more acuity gets pushed into skilled nursing facility environments in particular what happens to patients as they transition from that skilled nursing facility perhaps to home health, perhaps to outpatient, perhaps home with no other

service. What happens to them in that handoff is important.

And then I know the MAP and NQF have been doing work on home- and community-based service measures. I really would love to see more in that area, especially on home health.

Because again, under the bundling environment that CMS is putting forward under Medicare, comprehensive joints, now cardiac was just introduced, more and more patients are being not forced, but it's being highly recommended that patients be placed in lower level settings and home health is absorbing a lot of that acuity.

So really understanding what's happening as a result of that. While we're concentrated very much on the Medicare spending per beneficiary what are we really doing in terms of outcome.

CO-CHAIR SEPTIMUS: Thank you very much. Albert and then Pat.

MEMBER WU: This is a little potpourri of different things.

In the patient-reported outcome measures, if we were to supplement some of our surveys one question might be how incidence or errors are handled from the patient's point of view.

One question could be did people disclose what happened to you. Did they apologize. Were they honest. Were they empathetic.

I think -- maybe there's a gap perhaps that you're identifying.

A second sort of completely different dimension is we had some sort of sidebar discussion earlier about what NQF's standards are for what is acceptable reliability, what is a sufficient number of studies to provide evidence to support validity.

If you're doing a literature synthesis what's an adequate positive predictive value. So there are a number of things that I think that we've sort of got jotted down here and there.

But I'm not sure that we have thought

them through as clearly for performance measurement and performance improvement as opposed to sort of what is done in the meta analysis world. So that's -- codifying that across panels might be an interesting thing to do.

MR. LYZENGA: It's definitely something that's been considered before. We have had some work done by a testing task force, for example, and they I think put that question to them and they really resisted actually putting any thresholds, numbers for things like reliability and validity. But maybe it's something that we can revisit.

Something that there was discomfort with among that group at least and they really wanted to leave some space for committees to wrestle with these things in different scenarios, different measures, and different circumstances and come to a judgment.

But it is something that we've heard from our committee still that that's a really difficult thing. You know, committees want to

know what does this score mean, that 0.6, or 0.7, or 0.8 and what does that mean. Is that good, or is it bad, or is it somewhere in between and what's the cutoff.

We haven't really provided too definitive guidance around that yet. But again, maybe that's something that we could revisit in the future.

CO-CHAIR SEPTIMUS: Is someone by the way keeping track of these suggestions so that when we send out things? Okay. Very good. Pat.

MEMBER QUIGLEY: Thank you. For opportunities for measurement improvement I would like to suggest that there still be more efforts towards harmonization of these measures across the different organizations that are submitting them.

But also at some point in time to have a real critical analysis of how many of these are really value-added.

Because wherever I go I just hear the burden, all the burden that's associated with all these measures, and the data extraction, and the

reporting.

For opportunities for performance measures to have greatest impact on improving outcomes I would like to see more composite measures. And those who have composite measures to see if they've moved them to outcomes.

You know, we had one of the composite measures I think was NCQA, the group for ambulatory care. They had fall risk, screening, assessment and care planning, but have they moved to outcomes. Have they extended that composite measure to outcomes.

And still as people come forward for those that really could have the composite measure then if we had the structure and the process, if the outcomes weren't met then they could go back and look at structure and process.

And here might be an example of that. As you know, I was one of the dissenting votes for the PSI 90. And the measure that I had the most concern about of course was the one related to post surgical hip fractures.

And the AHRQ toolkit just came out two days ago. And I went and I looked into the toolkit because I would, a patient safety person.

And don't you know there's composite measures in the AHRQ toolkit.

Well, then I looked at the post-surgical fracture risk. And what AHRQ put out there as a best practice approach for prevention is standard fall prevention. For post surgical hip fractures.

It had nothing to do with assessing injury risk for the surgical population. It had nothing to do with identifying those at risk because of osteoporosis or prior hip fracture.

So, this has implications. So I think if we could really look at having composite measures. Because whatever comes out of here at some point is going to go towards implementation. So I would like to ask for that.

So I give that example as one. So I think harmonization.

But I also think at some point as NQF

goes forward there should be some expectations to stratify based on risk.

Some of these measures that come in have such a large age range. And they aren't stratifying based on risk, or age. So I think that there's some opportunity to still have more precision.

And that's what I ask for, is more precision. Because there's a lot of work that's being done by experts and sometimes we just don't get to have that.

So I know I get a little long-winded, but the last one for areas for investment for further measures to be submitted, I'd like to suggest that maybe early remobilization.

There's a lot of work being done in hospitals now for eliminating bedrest and getting people up and moving. And there is emerging evidence surrounding that. So I think early remobilization, decreasing bedrest, those kinds of things.

And then in terms of workforce safe

patient handling and movement. Preventing back injuries in the workforce I think would have a lot of strength. And there's enough evidence in that for that to come forward whether by the American Nurses Association, the organizations surrounding safe patient handling and mobility.

So those would be my comments and thanks for this opportunity.

CO-CHAIR SEPTIMUS: Thank you very much. Charlotte, did you have another one?

MEMBER ALEXANDER: I have one more. I'd like to see some more work around disparities. It's certainly something that I'm trying to get my hospital system more engaged with. I'm learning how poorly we identify patients and their languages and how can you communicate if you don't even know what language they're speaking.

And I think it's a huge patient safety issue. So if there's a way that we can come up with a measure that can help us identify even if it's just language, much less all the other stuff that we need to be looking at, I'd like to see that.

MR. LYZENGA: I know we do have a committee on disparities and that's a major focus of NQF, trying to figure those things out.

I also know there are a lot of problems trying to get the data to identify disparities comprehensively and accurately and kind of nail that stuff down. So it's something that there's a lot of people thinking about and it's very important.

MEMBER MCGIFFERT: Since we're talking about patient safety are you specifically thinking of disparities with patient safety, or just in general? Because that's kind of -- in those discussions there's been a pretty consistent philosophy that we don't include patient safety issues.

MEMBER ALEXANDER: I think we have to. I mean, whether it is dialysis rates with black men, whether it is a language issue that we're not communicating with the patient in their own language, whether it's pain medicine that black people are getting in the emergency room. There's

so much in the way of safety and quality disparities that exist.

MEMBER MCGIFFERT: Like medication.

(Simultaneous speaking)

MEMBER MCGIFFERT: But it wouldn't be like a CLABSI --

MEMBER ALEXANDER: Access to heart disease, access to cancer care.

MEMBER MCGIFFERT: A CLABSI and a surgical error you think, maybe?

MEMBER ALEXANDER: I think that when you look at people that, for instance, are not English speaking or have limited English proficiency their complication rate increases in the hospital. Their length of stay increases a little bit. So, there are definite quality aspects that are tied to it.

We're stratifying and that's great, but we're really not getting to the root of going the next step to try and force people to fix it. So I'd like to see a little more work in that area.

CO-CHAIR SEPTIMUS: Well, we're just

at about the top of the hour and we've got some great suggestions.

Unless staff has anything else to say I want to say this is a phenomenal meeting as always. Obviously our portfolio this time in terms of time and numbers were much more manageable which is why we still have a quorum and we're finishing seven minutes before the time of ending.

But really, if it wasn't for the great work that everybody does beforehand in preparing to present these measures, and thank for the developers of course for all the time that they do, and of course as I mentioned before the incredible heavy lifting that the NQF staff does which we couldn't even begin to start if it wasn't for the work that they do before we get here. So obviously a big thanks to them.

And with that I wish you all a great rest of the summer. And from Iona and I thank you so much.

(Whereupon, the above-entitled matter went off the record at 2:52 p.m.)